

**Министерство цифрового развития, связи и массовых коммуникаций
Российской Федерации ордена Трудового Красного Знамени федеральное
государственное бюджетное образовательное учреждение
высшего образования
«Московский технический университет связи и информатики»
(МТУСИ)**

На правах рукописи



A handwritten signature in blue ink, appearing to be 'Lat', is located to the right of the MTSU logo.

Харрасов Камиль Раисович

**Разработка нейросетевых методов обнаружения дорожно-транспортных
происшествий на основе обработки визуальной информации**

Специальность 2.3.8.
«Информатика и информационные процессы»

Диссертация на соискание учёной степени
кандидата технических наук

Научный руководитель:
кандидат технических наук, доцент
Мосева Марина Сергеевна

Москва – 2026

Введение 4

Глава 1. Анализ существующих методов и алгоритмов решения задачи обработки изображений	12
1.1. Анализ современных подходов к распознаванию дорожно-транспортных происшествий в условиях ограниченной видимости.....	13
1.2. Модель угроз возникновения и классификации дорожно-транспортных происшествий.....	15
1.3. Подходы к обнаружению ДТП	17
1.4. Проблемы визуального распознавания ДТП в условиях затрудненной видимости	24
1.5. Методы борьбы с зашумлениями.....	26
1.6. Выводы по главе	34
Глава 2. Исследование и разработка метода определения дорожно-транспортных происшествий.....	36
2.1. Исследование методов определения ДТП.....	37
2.2. Разработка метода обнаружения ДТП	42
2.3. Выводы по главе	52
Глава 3. Метод восстановления изображения от погодных зашумлений ...	54
3.1. Сравнительный анализ методов удаления дождя с изображений	54
3.2. Исследование базовых методов	58
3.3. Разработка метода отчистки изображения от погодных шумов	65
3.4. Выводы по главе	74
Глава 4. Разработка АИС для обнаружения ДТП в затрудненных погодных условиях.....	75
4.1. Архитектура АИС	76

4.2.	Функциональная структура АИС	78
4.3.	Организация взаимодействия процессов в АИС	84
4.4.	Информационная модель АИС.....	87
4.5.	Апробация и оценка эффективности разработанной АИС.....	91
4.6.	Выводы по главе	99
Заключение		101
Список литературы		103
Приложение А. Модель угроз возникновения дорожно-транспортных происшествий		117
Приложение Б. Свидетельства о государственной регистрации программ для ЭВМ.....		122
Приложение В. Акты о внедрении		126

Введение

Современные интеллектуальные системы видеонаблюдения и автоматического анализа дорожной обстановки становятся неотъемлемой частью инфраструктуры «умного» транспорта, систем помощи водителю (ADAS) и городских центров управления дорожным движением. Одной из ключевых задач таких систем является автоматическое выявление дорожно-транспортных происшествий (ДТП) в режиме реального времени, что позволяет оперативно реагировать на инциденты, снижать последствия аварий и повышать общую безопасность на дорогах.

Однако одной из значительных проблем, существенно влияющих на качество распознавания ДТП, является ухудшение видимости вследствие погодных условий: тумана, дождя, снега, засветки, загрязнения оптики и других факторов. Наличие атмосферных шумов приводит к искажению визуальной информации, снижает качество изображений и, как следствие, ухудшает эффективность работы систем компьютерного зрения. Повышение эффективности предотвращения и распознавания ДТП требует внедрения алгоритмов, способных функционировать при внешних погодных факторах. Системы компьютерного зрения становятся ключевым инструментом своевременного автоматизированного обнаружения событий, однако они требуют дальнейших исследований и адаптации к условиям реальной городской среды.

Степень разработанности темы исследования. По тематике диссертационного исследования опубликовано значительное количество работ российских и зарубежных ученых. Среди российских исследователей следует отметить М. В. Терешонка, В. В. Зиядинова, И. С. Кружилова, О. Ю. Шамаеву, А. С. Катасёва. Среди зарубежных исследователей значительный вклад в развитие соответствующих научных направлений внесли И. Гудфеллоу, С. К. Найар, С. Г. Нарасимхан, Р. Т. Тан, В. Янг, С. В. Замир, В. М. Патель, Я. Яо, К.

Павар, В. Аттар. Их работы охватывают различные аспекты нейросетевой обработки и распознавания визуальной информации, анализа дорожных сцен и динамики транспортных средств, обнаружения дорожно-транспортных происшествий, восстановления изображений при атмосферных искажениях и построения вычислительно эффективных моделей компьютерного зрения. Полученные ими результаты сформировали научно-методическую основу для развития методов обработки визуальной информации, применяемых при решении задач обнаружения дорожно-транспортных происшествий в условиях затрудненной видимости. Несмотря на достигнутые результаты, применительно к автоматизированному обнаружению дорожно-транспортных происшествий по данным городского видеонаблюдения сохраняются следующие направления дальнейших исследований:

- **Совместный анализ динамических и визуальных признаков дорожной сцены:** методы, основанные только на покадровой классификации, не учитывают в полной мере динамику развития аварийной ситуации, тогда как методы анализа траекторий зависят от устойчивости локализации и сопровождения транспортных средств.
- **Восстановление визуальной информации при атмосферных искажениях с учётом задач последующего анализа:** для обработки видеопотоков существенное значение имеет не только визуальное качество восстановленного изображения, но и сохранение признаков, используемых алгоритмами детекции объектов.

Целью диссертационного исследования является повышение эффективности обнаружения дорожно-транспортных происшествий в условиях затрудненной видимости посредством разработки нейросетевых методов обработки визуальной информации.

Для достижения цели сформулированы следующие **задачи**:

- Исследовать существующие методы обнаружения дорожно-транспортных происшествий и методы восстановления визуальной информации в условиях затрудненной видимости.

- Разработать нейросетевой метод обнаружения дорожно-транспортных происшествий посредством анализа видеопотока с уличных городских камер.
- Разработать нейросетевой метод восстановления изображения от зашумлений, вызванных атмосферными осадками.
- Разработать архитектуру автоматизированной информационной системы определения дорожно-транспортных происшествий в условиях затрудненной видимости, реализующую разработанные методы.

Объектом исследования являются информационные процессы цифровой обработки визуальной информации, а также автоматизированные информационные системы, предназначенные для обнаружения событий.

Предметом исследования являются методы, алгоритмы и программно-технические средства цифровой обработки визуальной информации, обеспечивающие восстановление изображений, анализ дорожной сцены и автоматизированное обнаружение дорожно-транспортных происшествий.

Методы исследования. Для решения сформулированных задач использовались методы компьютерного зрения и глубокого обучения, автоматической обработки визуальной информации, методы статистического анализа экспериментальных данных.

Научная новизна результатов диссертации заключается в совокупности взаимосвязанных программных и технических решений для автоматизации распознавания дорожно-транспортных происшествий в условиях затрудненной видимости.

В процессе выполнения диссертационной работы получены следующие оригинальные научные результаты:

1. Разработан нейросетевой метод цифровой обработки визуальной информации, *отличающийся* двухэтапным принятием решения об обнаружении происшествия, включающим формирование событий-кандидатов по результатам анализа динамики транспортных средств и

их последующую нейросетевую классификацию, *позволяющий* учитывать динамику развития дорожной сцены при анализе видеопоследовательности (2.3.8, п.4).

2. Разработан нейросетевой метод восстановления изображения от зашумлений, вызванных атмосферными осадками, *отличающийся* сниженной параметрической ёмкостью генератора, *позволяющий* сохранять информативность визуальных признаков для последующего обнаружения объектов дорожной сцены с двукратным сокращением времени обработки (2.3.8, п.4).
3. Разработана архитектура автоматизированной информационной системы обнаружения дорожно-транспортных происшествий по визуальной информации, *отличающаяся* совместным применением средств восстановления информативности изображений дорожной сцены в условиях атмосферных осадков и двухэтапного механизма распознавания дорожно-транспортных происшествий, *позволяющая* адаптивно формировать последовательность выполнения процедур мониторинга, восстановления изображений и обнаружения происшествий (2.3.8, п.9).

Теоретическая и практическая значимость исследования определяется развитием методов цифровой обработки визуальной информации для обнаружения дорожно-транспортных происшествий в условиях затрудненной видимости. Полученные результаты расширяют возможности анализа видеопоследовательностей за счёт объединения динамических характеристик транспортных средств и визуальных признаков дорожной сцены, а также развивают методы восстановления изображений с учётом сохранения признаков, значимых для последующего анализа изображения. Разработанные методы создают основу для комплексной обработки визуальной информации при распознавании событий в изменяющихся условиях наблюдения. Практическая значимость заключается в возможности использования разработанных методов и архитектуры автоматизированной информационной системы в

интеллектуальных транспортных системах на основе существующей инфраструктуры городского видеонаблюдения.

Положения, выносимые на защиту

1. Метод цифровой обработки визуальной информации дорожной сцены, *позволяющий* обеспечивать полноту выявления дорожно-транспортных происшествий не ниже 0,88 на независимой тестовой выборке видеопоследовательностей.
2. Метод восстановления изображения от зашумлений, вызванных атмосферными осадками, *позволяющий* сохранить структурное сходство восстановленного изображения не менее 0,78 и информативность визуальных признаков для последующей детекции объектов с двукратным сокращением времени обработки кадра.
3. Архитектура автоматизированной информационной системы обнаружения дорожно-транспортных происшествий, *позволяющая* обеспечивать последовательную обработку с адаптивным включением восстановления изображения при изменении условий наблюдения и масштабируемое взаимодействие функциональных компонентов в зависимости от количества видеопотоков.

Степень достоверности и апробация результатов работы определяется использованием современных методов детектирования и классификации объектов и событий на изображениях, а также экспериментальной проверкой разработанных решений. Методическую основу исследования составляют методы анализа данных, машинного и глубокого обучения и математической статистики, адаптированные к решению поставленных задач.

Результаты были представлены на ряде международных конференций:

- 2024 Systems of Signals Generating and Processing in the Field of on Board Communications;
- 2024 Wave Electronics and Its Application in Information and Telecommunication Systems;

- 2024 Systems of Signal Synchronization, Generating and Processing in Telecommunications;
- 2025 Systems of Signals Generating and Processing in the Field of on Board Communications;
- 2026 Systems of Signal Synchronization, Generating and Processing in Telecommunications.

Результаты также обсуждались на заседании кафедры, а также на научных семинарах в МТУСИ.

Апробация результатов подтверждается свидетельствами о регистрации программ для ЭВМ, представленными в приложении А.

Личный вклад.

В ходе исследования автором лично были разработаны и реализованы ключевые подходы, направленные на повышение эффективности обнаружения дорожно-транспортных происшествий:

- Разработан нейросетевой метод цифровой обработки визуальной информации для обнаружения дорожно-транспортных происшествий, реализующий двухэтапное принятие решения на основе анализа динамики транспортных средств и последующей нейросетевой классификации событий-кандидатов, обеспечивающий полноту выявления дорожно-транспортных происшествий не ниже 0,88 при времени обработки кадра не более 20 мс.
- Разработан нейросетевой метод восстановления изображений от зашумлений, вызванных атмосферными осадками, основанный на применении облегчённого генератора с уменьшенной параметрической ёмкостью; обеспечивающих сохранение структурного сходства восстановленного изображения не ниже 0,78 при времени обработки кадра не более 29 мс, что соответствует двукратному сокращению времени обработки по сравнению с базовым методом.
- Разработана архитектура автоматизированной информационной системы обнаружения дорожно-транспортных происшествий,

обеспечивающая совместное применение разработанных методов, адаптивное включение восстановления изображения при изменении условий наблюдения и масштабирование вычислительных компонентов в зависимости от количества входных видеопотоков.

- Организованы и проведены экспериментальные исследования разработанных методов и архитектуры АИС, включающие оценку обнаружения дорожно-транспортных происшествий на независимой тестовой выборке видеопоследовательностей, исследование качества восстановления изображений и последующей детекции объектов, апробацию полного конвейера обработки и нагрузочное тестирование при различном количестве входных видеопотоков.

Реализация и внедрение. Алгоритмы и архитектура программно-аппаратного комплекса, разработанные в настоящей работе, внедрены в следующих организациях:

- ООО «Интелсити»
- ООО «ТАГИР»

Также разработанное алгоритмическое обеспечение внедрено в учебный процесс кафедры «Математическая кибернетика и информационные технологии» МТУСИ. Подтверждается соответствующими актами внедрения результатов диссертационной работы.

Соответствие специальности. Тематика и результаты диссертации соответствуют следующим направлениям специальности: 2.3.8. «Информатика и информационные процессы».

- п.4 – разработка методов и технологий цифровой обработки аудиовизуальной информации с целью обнаружения закономерностей в данных, включая обработку текстовых и иных изображений, видео контента. Разработка методов и моделей распознавания, понимания и синтеза речи, принципов и методов извлечения требуемой информации из текстов.

- п.9 – разработка архитектур программно-аппаратных комплексов поддержки цифровых технологий сбора, хранения и передачи информации в инфокоммуникационных системах, в том числе, с использованием «облачных» интернет-технологий и оценка их эффективности.

Публикации. По материалам диссертации опубликовано 18 научных работ, из них 14 научных статей, из которых 3 в научных журналах и изданиях, рекомендованных ВАК при Министерстве науки и высшего образования Российской Федерации, в периодических научных журналах индексируемых Scopus – 11, в том числе 1 работа опубликована без соавторов, получены 4 свидетельства о регистрации программы для ЭВМ.

Объем и структура работы. Диссертация состоит из введения, 4 глав, заключения и 3 приложений. Полный объем диссертации составляет 128 страниц, включая 14 рисунков и 14 таблиц. Список литературы содержит 109 наименований.

Глава 1. Анализ существующих методов и алгоритмов решения задачи обработки изображений

В первой главе обоснована необходимость разработки методов обнаружения дорожно-транспортных происшествий на основе анализа визуальной информации в условиях городской среды. Показано, что существующие методы, основанные на использовании сенсорных данных и алгоритмах компьютерного зрения, обладают рядом ограничений, связанных с зависимостью от условий наблюдения, качеством входных данных и вычислительными требованиями.

Проведён анализ методов обнаружения дорожно-транспортных происшествий. Установлено, что атмосферные искажения ухудшают качество визуальных признаков дорожной сцены и снижают устойчивость методов обнаружения, что отражается на результатах обнаружения ДТП.

Рассмотрены существующие методы восстановления изображений, включая классические алгоритмы и подходы на основе нейронных сетей. Несмотря на возможность повышения визуального качества, их использование при обработке видеопотоков связано с ограничениями по вычислительной сложности и необходимостью сохранения структурных признаков, используемых при последующем анализе изображений

По результатам проведенного анализа сформулированы требования к методам обнаружения ДТП и восстановления изображений, ориентированные на обработку визуальных данных при ухудшении их качества и возможность функционирования в режиме реального времени.

1.1. Анализ современных подходов к распознаванию дорожно-транспортных происшествий в условиях ограниченной видимости

Внедрение интеллектуальных систем и технологий машинного зрения в разные области жизнедеятельности человека привело к значительным изменениям в подходах к обеспечению безопасности. Несмотря на значительные успехи в применении интеллектуальных систем, существующие решения в области мониторинга дорожного движения по-прежнему не обеспечивают необходимого уровня надежности. Рост числа транспортных средств, расширение дорожной инфраструктуры и увеличение плотности городского дорожного движения сопровождаются как положительными изменениями в сфере общественной мобильности, так и серьезными вызовами для безопасности дорожного движения. По данным Национального центра безопасности дорожного движения Министерства внутренних дел России, в 2025 году было зарегистрировано 64,07 млн транспортных средств — на 1,5 % больше, чем в предыдущем году [1]. Особенно быстрый рост наблюдается в крупных городах, где плотность транспортных средств на единицу уличной и дорожной сети достигает критических значений.

Увеличение транспортной насыщенности не может не сказываться на аварийности: несмотря на внедрение интеллектуальных систем управления движением и модернизацию инфраструктуры, дорожно-транспортные происшествия (ДТП) остаются одной из основных причин гибели и травматизма в мирное время. Согласно ПДД РФ (раздел 1, общие положения), ДТП – событие, возникшее в процессе движения по дороге транспортного средства и с его участием, при котором погибли или ранены люди, повреждены транспортные средства, сооружения, грузы либо причинен иной материальный ущерб.

Ежегодно, по данным Всемирной организации здравоохранения (ВОЗ), в результате ДТП погибает около 1,35 миллиона человек, а еще от 20 до 50

миллионов получают травмы различной степени тяжести. Это оказывает значительное давление на системы здравоохранения и экономики как в развитых, так и в развивающихся странах. Согласно статистике ГИБДД, ежегодно в стране фиксируется порядка 120-150 тысяч ДТП с пострадавшими. При этом, по данным Всемирной организации здравоохранения, Россия входит в число стран с высоким уровнем смертности на дорогах: более 13 погибших на 100 тысяч человек населения.

Помимо угрозы жизни и здоровью граждан, ДТП имеют и масштабные косвенные последствия. Каждое происшествие в условиях городской среды — это локальный сбой в работе всей транспортной системы. Возникают заторы, ограничивается проезд для общественного транспорта и экстренных служб, нарушаются логистические цепочки. Одно несвоевременно устраненное ДТП в час пик может привести к масштабным "хвостам" в пробках, затрагивая районы, находящиеся в десятках километров от места инцидента. К этому добавляются экономические потери: простой транспорта, перерасход топлива, несоблюдение графиков доставки, снижение пропускной способности дорожной сети. По оценкам ВОЗ, экономический ущерб от ДТП может достигать 2–3% ВВП страны ежегодно.

Экономические потери от ДТП включают:

- **Прямые расходы.** Затраты на медицинскую помощь, восстановление транспортных средств и дорожной инфраструктуры;
- **Косвенные убытки.** Потеря трудоспособности, снижение производительности труда, транспортные заторы и замедление логистических процессов;
- **Инфраструктурные издержки.** Необходимость реорганизации движения, модернизации систем безопасности и видеонаблюдения.

По оценкам ВОЗ, совокупные издержки, связанные с дорожно-транспортными происшествиями, могут достигать 3% ВВП страны, особенно в регионах с недостаточно развитыми системами мониторинга и реагирования на ДТП.

Согласно данным Государственной инспекции безопасности дорожного движения (ГИБДД), в период с 2015 по 2025 год в России количество дорожно-транспортных происшествий сократилось на 37%, число пострадавших – на 37%, а погибших – на 44%. Вместе с тем соотношение числа пострадавших и погибших к общему количеству ДТП за рассматриваемый период практически не изменилось, что указывает на сохранение тяжести последствий дорожных происшествий.

В условиях роста числа транспортных средств и высокой интенсивности движения в городах особое значение приобретает своевременное обнаружение ДТП. Сокращение времени между возникновением происшествия и его фиксацией позволяет ускорить реагирование экстренных служб, повысить вероятность своевременного оказания помощи пострадавшим, а также уменьшить риск возникновения вторичных происшествий и продолжительных транспортных заторов. Кроме того, оперативное выявление аварий способствует снижению связанных с ними экономических и инфраструктурных издержек. Таким образом, актуальной является разработка интеллектуальных систем мониторинга, обеспечивающих автоматическое обнаружение ДТП в режиме реального времени, в том числе при затрудненной видимости и наличии визуальных искажений

1.2. Модель угроз возникновения и классификации дорожно-транспортных происшествий

ДТП обусловлены множеством разнообразных факторов – от погодных условий и состояния дорог до ошибок водителей и механических неисправностей. Эти факторы действуют как по отдельности, так и в совокупности, приводя к возникновению широкого спектра сценариев

возникновения и развития ДТП. Комплексный характер событий требует не только статистического анализа, но и системного подхода к описанию условий.

Для правильной разработки методов обнаружения дорожно-транспортных происшествий важно учитывать не только тип происшествия, но и контекст, в котором оно происходит – где и когда они происходят, а также скорость, с которой развиваются события.

Анализ статистических данных показывает, что значительная доля дорожно-транспортных происшествий происходит:

- на перекрестках и развязках, где пересекаются транспортные потоки, в том числе при повороте налево и выезде с второстепенной дороги [2];
- вблизи пешеходных переходов и остановок общественного транспорта, особенно в ночное время [1];
- на загородных автомагистралях, где преобладают ДТП, связанные с высокой скоростью [1];
- в городских районах с интенсивным движением, особенно в часы пик, когда количество конфликтных ситуаций увеличивается [1].

Что касается временных факторов, при анализе рисков ДТП следует учитывать следующее:

- темное время суток, особенно в районах с недостаточным освещением: в 2025 году на эту категорию пришлось 32,5 % ДТП, повлекших за собой гибель или травмы людей [1];
- зимний и осенний сезоны, когда осадки, гололед и туман могут ухудшать дорожные условия и видимость [2];
- вечерние часы: в 2025 году 34,1 % ДТП, повлекших за собой гибель или телесные повреждения, произошло в период с 16:00 до 21:00 [1].

На основе результатов анализа литературы, нормативно-правовой базы и статистических исследований модель риска дорожно-транспортных происшествий описана в Приложении А для дальнейшего анализа типов происшествий, условий их возникновения и масштабов последствий с целью

определения наиболее значимых сценариев ДТП и построения соответствующих систем обнаружения.

1.3. Подходы к обнаружению ДТП

Современные системы видеонаблюдения и интеллектуального анализа базируются на различных типах сенсоров и архитектурных решений – от простых видеокамер до комплексных мультимодальных платформ, объединяющих данные от радаров, GPS, акселерометров и облачных источников [3]. Технические ограничения и возможности этих систем напрямую влияют на точность, скорость и устойчивость алгоритмов распознавания.

Подход, основанный на датчиках. Аппаратный подход определения ДТП основан на использовании разнообразных датчиков и сенсоров, размещенных на дорожной инфраструктуре и в транспортных средствах. Эти устройства измеряют большое количество различных параметров, например, скорость движения транспортного средства, средняя скорость движения потока транспортных средств, расстояние между транспортными средствами и многие другие параметры. При обнаружении неполадок или аномалий в информации, полученной с аппаратного обеспечения (датчиков, сенсоров и т.п.) срабатывает предназначенная для анализа этой информации система, что в свою очередь сигнализирует о возможном ДТП.

Для автоматического распознавания ДТП применяются разнообразные бортовые сенсоры: инерциальные, позиционные, оптические, акустические и др. Акселерометры регистрируют резкие перегрузки и вибрации во время столкновений [4], гироскопы фиксируют изменения углового положения, а модуль GPS позволяет определять координаты места ДТП и траекторию движения до столкновения [4]. В отличие от предыдущей группы датчиков, видеокамеры фиксируют визуальные признаки ДТП – от появления препятствия

до момента столкновения, – которые затем могут быть использованы для автоматической классификации событий или для анализа обстоятельств ДТП специалистами [5]. Радарные дальномеры и лидары представляют собой еще одну категорию бортовых датчиков, которые активно излучают сигнал и измеряют расстояние до объектов на основе его отражения для обнаружения ДТП. Они способны заранее обнаруживать препятствия, находящиеся впереди транспортного средства, и тем самым либо предупреждать водителя, либо инициировать превентивное торможение, снижая таким образом скорость столкновения. Особо следует отметить датчики, контролирующие состояние водителя и влияние человеческого фактора на безопасность участников дорожного движения: например, датчики, регистрирующие моргание глаз или концентрацию алкоголя в выдыхаемом воздухе. Эти устройства не указывают напрямую на факт аварии, но позволяют выявить факторы риска (сонливость, опьянение) и тем самым предотвратить потерю управления еще до ДТП [6, 7]. Используемые в транспортных средствах датчики можно условно разделить на пассивные и активные. Пассивные (акселерометры, гироскопы, видеокамеры, микрофоны) воспринимают изменения состояния или внешней среды без излучения сигнала, тогда как активные (радары, ультразвуковые дальномеры, лидары) излучают собственный сигнал и улавливают его отражения. Современные автомобили нередко оснащены комбинацией обоих типов сенсоров в составе продвинутых систем помощи водителю, что создает поток данных для выявления аварийных событий.

Помимо штатных датчиков автомобиля, в роли аварийных детекторов могут выступать и внешние по отношению к машине устройства, которые находятся непосредственно в транспорте. В исследованиях встречается применение смартфонов, оборудованных акселерометрами, гироскопами, GPS и микрофонами – эти встроенные в телефон сенсоры могут фиксировать характерные признаки ДТП (удар, резкое замедление, звук столкновения) даже в тех случаях, когда сам автомобиль не оснащен специальной электроникой.

Аналогично, для мотоциклистов перспективными являются умные защитные шлемы с интегрированными датчиками удара и положения [8].

Бортовые системы автоматического обнаружения аварий работают автономно. Они отслеживают показания датчиков и при распознавании столкновения непосредственно из салона инициируют аварийный вызов. При срабатывании в машине датчика удара специальный модуль автоматически осуществляет голосовой вызов экстренной службы и передает ей минимальный набор данных о происшествии. Аналогичные службы используют бортовые телематические устройства, которые связываются с оператором и сообщают координаты аварии. Преимущество локального размещения – независимость от внешних устройств: даже если авария произошла вдали от других участников движения, автомобиль сам распознает ее и пытается вызвать помощь.

Подход, основанный на анализе видеопотока. Системы обнаружения аномалий на дорогах, использующие видеопотоки с камер наблюдения за дорожным движением, различаются по способу представления наблюдаемого события. Один из подходов заключается в анализе изменений дорожной обстановки во времени и включает отслеживание транспортных средств по последовательности кадров. Другой подход основан преимущественно на визуальных признаках отдельных изображений или ограниченном временном контексте.

При анализе динамики дорожной обстановки дорожно-транспортное происшествие рассматривается как результат изменений в поведении участников дорожного движения во времени. В данном случае информативными признаками являются пространственно-временные характеристики, описывающие траектории движения транспортных средств, изменения их скорости и направления движения, их относительные положения, а также характер их взаимодействий.

Метод на основе нейронных сетей для обнаружения и прогнозирования дорожно-транспортных происшествий, представленный в [9], основан на интеграции визуальной информации и контекста дорожной обстановки.

Признаки зарождающегося ДТП выявляются путем совместного анализа информации о транспортных средствах и последовательности кадров. Однако использованные видеоданные были получены с видеорегистратора, установленного на движущемся транспортном средстве, что ограничивает возможность прямого переноса данного метода на стационарные системы городского видеонаблюдения.

В работе [10] лежит иной принцип: после локализации транспортных средств с помощью нейронной сети осуществляется отслеживание их центроидов. Сформированные таким образом траектории позволяют обнаруживать движение против потока транспорта, резкие изменения направления и другие отклонения, характерные для потенциально опасных ситуаций. Возможность связать обнаруженную аномалию с изменением движения конкретного транспортного средства обеспечивает интерпретируемость результатов; однако ошибки обнаружения, вызванные заслонением объектов и плотным транспортным потоком, напрямую влияют на сгенерированные траектории.

В работе [11] также используются временные изменения пространственных характеристик дорожной сцены. Из последовательности кадров извлекаются признаки движения, что позволяет обнаруживать резкие изменения скорости и относительных положений транспортных средств. В данном случае обнаружение столкновений определяется характером развития события; это требует стабильной сегментации объектов на протяжении всей последовательности изображений, что повышает чувствительность метода к ухудшению качества видео.

Еще один метод выявления аварийных ситуаций представлен в [12]. Обучение проводится на сценариях нормального дорожного движения, после чего аномальное событие идентифицируется на основе ошибки в прогнозировании развития сцены или реконструкции видеопоследовательности. Этот принцип позволяет обнаруживать ранее ненаблюдаемые типы инцидентов без необходимости предварительной ручной аннотации. Однако при переносе

полученных результатов на стационарные камеры наблюдения за дорожным движением необходимо также учитывать использование видеоданных, полученных с борта транспортного средства.

В исследовании [13] анализ аварийных ситуаций выполняется с учетом взаимодействия участников дорожного движения на последовательности кадров. Для выявления опасных сближений и столкновений оцениваются изменения расстояния и взаимного положения транспортных средств. Использование временной последовательности позволяет учитывать развитие дорожной ситуации, однако требует дополнительных вычислительных ресурсов и устойчивого сопровождения объектов.

Анализ динамики дорожной обстановки позволяет охарактеризовать дорожно-транспортное происшествие не только по его визуальным последствиям, но и по изменениям в поведении участников дорожного движения, предшествующим столкновению или сопровождающим его. В то же время методы, использующие траектории и другие пространственно-временные характеристики транспортных средств, зависят от надежности их локализации и отслеживания. При ухудшении видимости ошибки, возникающие на этих этапах, могут повлиять на последующий анализ развития дорожной ситуации.

В отличие от динамического анализа, методы, обрабатывающие отдельные кадры, идентифицируют дорожно-транспортные происшествия в первую очередь на основе визуального состояния дорожной сцены. В этом случае решение принимается на основе пространственных характеристик текущего изображения или небольшого числа последовательных кадров, без необходимости длительного отслеживания транспортных средств.

Для классификации дорожных сцен в [14] используется модифицированная архитектура Inception-V4. Предварительная сегментация видеопотока выполняется с использованием индекса структурного сходства, что позволяет учитывать изменения между последовательными кадрами без непосредственного построения траекторий объектов. В [15] для извлечения признаков используются EfficientNet [16] и MobileNetV2 [17], а недостаток

реальных данных о дорожно-транспортных происшествиях компенсируется включением в процесс обучения синтетических изображений сценариев ДТП.

В [18] предложен специфический подход к обработке видеопотока. Кадры сначала помещаются в буфер, после чего классифицируются с помощью сверточной нейронной сети. Запись видео инициируется при обнаружении события, что позволяет снизить вычислительную нагрузку и объем хранимой информации.

Отдельное направление представляют методы, рассматривающие ДТП как аномальное изменение дорожной сцены. В [19] для выявления таких изменений применяются сверточный автокодер и LSTM, обучение которых проводится на данных, характеризующих нормальное движение. Потенциальное ДТП фиксируется при отклонении наблюдаемой видеопоследовательности от сформированных пространственно-временных закономерностей. Сходный подход используется в [20], где нормальное состояние сцены описывается с помощью группы шумоподавляющих автокодеров, а признаком аномального события служат рост ошибки реконструкции и изменение взаимного расположения транспортных средств.

Помимо прямой классификации и обнаружения аномалий, используются также комбинированные решения. В [21] классификация состояния сцены дополняется предварительным обнаружением и отслеживанием объектов, а для анализа зоны потенциального столкновения используется дескриптор интенсивности дорожного движения. В методе, описанном в [22], также используется обнаружение транспортных средств на основе CNN, однако дальнейший анализ основан на отслеживании транспортных средств и оценке их скорости. В [23] траектории используются для обнаружения внезапных изменений в характере движения и возникающих столкновений в режиме реального времени.

Рассмотренные методы анализа видеоданных различаются по характеру признаков, используемых для обнаружения ДТП. К ним относятся визуальные характеристики дорожной сцены, аномальные отклонения от нормального

движения, а также параметры движения и взаимодействия транспортных средств. Анализ отдельных кадров снижает зависимость от устойчивости длительного сопровождения объектов, но предъявляет требования к сохранности информативных признаков изображения. Учет пространственно-временных характеристик, напротив, дает возможность анализировать развитие дорожной ситуации во времени, при этом результат в большей степени зависит от качества локализации и последующего сопровождения транспортных средств.

Преимущества и ограничения рассмотренных подходов.

Существующие подходы к автоматическому обнаружению ДТП различаются по источникам входных данных, масштабу и уровню наблюдаемой дорожной сцены и требованиям к инфраструктуре для эксплуатации. Системы, основанные на бортовых датчиках транспортных средств или дорожной инфраструктуре, обеспечивают высокую точность регистрации физических параметров, сопровождающих столкновения, включая ускорения и изменения ориентации транспортного средства. Однако их применение связано с необходимостью оснащения автомобилей дополнительным оборудованием либо широкого покрытия сенсорными комплексами УДС.

Развитая инфраструктура видеонаблюдения на УДС создает условия для масштабирования методов обнаружения ДТП без необходимости дополнительного оснащения транспортных средств аппаратными средствами. По сравнению с сенсорными системами видеоданные содержат информацию не только об отдельных участниках движения, но и об окружающей дорожной обстановке, что позволяет анализировать их взаимодействие и пространственные характеристики происшествия.

Рассмотренные методы по-разному используют визуальную информацию при обнаружении ДТП. Анализ видеопотока позволяет учитывать изменение состояния дорожной сцены и взаимодействие участников движения во времени, однако результат зависит от устойчивости детектирования и сопровождения объектов [24]. При классификации отдельных кадров такая зависимость выражена в меньшей степени, но ухудшение качества изображения приводит к

деградации визуальных признаков, используемых для распознавания дорожной ситуации.

1.4. Проблемы визуального распознавания ДТП в условиях затрудненной видимости

Применение систем компьютерного зрения в условиях уличной эксплуатации приводит к получению искажений изображения, связанных как с атмосферными явлениями, так и с эксплуатационными факторами [25]. Атмосферные искажения ухудшают различимость контуров транспортных средств, снижают надежность их обнаружения в последовательности кадров и нарушают временную согласованность наблюдений, затрудняя выявление признаков возникающей чрезвычайной ситуации. В работах [26, 27] было исследовано влияние шума на эффективность сверточных нейронных сетей и установлена зависимость между качеством распознавания и степенью искажения визуальных данных.

Туман и дымка являются одними из наиболее изученных атмосферных факторов, приводящих к ухудшению качества изображений [28]. В соответствии с классической моделью атмосферного рассеяния [29] наблюдаемое изображение формируется под воздействием ослабления прямого светового потока и рассеянного атмосферного света (airlight). Это приводит к снижению локального контраста, размыванию границ объектов и изменению цветопередачи. Как показано в [30], увеличение концентрации аэрозольных частиц сопровождается экспоненциальным сокращением дальности видимости и снижением точности нейросетевых методов обнаружения объектов при увеличении оптической плотности среды

Дождь характеризуется более сложным характером визуальных искажений. Следы движущихся капель образуют вертикальные и наклонные

полосы, создающие дополнительные градиентные структуры и затрудняющие корректное выделение границ объектов [31]. Мелкодисперсные капли формируют эффект «вуали» и снижают контрастность изображения, тогда как крупные капли, попадающие на объектив или ветровое стекло, способны частично перекрывать наблюдаемую сцену. В [32, 33] показано, что дождь вносит как аддитивные, так и мультипликативные шумовые составляющие, отрицательно влияющие на результаты детектирования объектов

Снег действует аналогично дождю, но частицы отличаются большим размером и разнообразием формы, из-за чего шум носит нерегулярный и динамический характер [34, 35]. Он сложное и многоаспектное воздействие на качество изображения и эффективность компьютерного зрения. Частицы осадков формируют нерегулярные текстурные шумы – от быстро движущихся снежинок до оседающего покрытия на поверхности – что существенно деформирует фон, контуры объектов и изменяет их отражательные свойства. В работе по симуляции снегопада [36] показано, как метод геометрического рассеяния снежного шума, наложенного на 3D-облако точек, позволяет оценить и компенсировать ухудшение точности детекции движущихся объектов, особенно во влажных условиях покрытия дороги. В мультимодальном подходе [37] описаны четкие параметры деградации: артефакты снега приводят к искажениям изображения и повсеместному шуму, показывающие снижение точности детекции объектов. В исследовании [38] установлено, что снег вызывает размывание форм и скрытие границ – а это приводит к падению активаций и перестройке карты признаков, вызывая падение точности детекции до 70 %.

Атмосферные искажения ухудшают различимость контуров транспортных средств, снижают надежность их обнаружения в последовательности кадров и нарушают временную согласованность наблюдений. В работах [26, 27] было исследовано влияние шума на эффективность сверточных нейронных сетей и установлена зависимость между качеством распознавания и степенью искажения визуальных данных. В неблагоприятных погодных условиях ухудшение

качества визуальной информации затрудняет выявление признаков возникающей чрезвычайной ситуации, включая изменения относительных положений транспортных средств и характера их движения. Это снижает эффективность обнаружения дорожно-транспортных происшествий с помощью методов, основанных на динамических характеристиках дорожной обстановки и взаимодействиях между участниками дорожного движения.

1.5. Методы борьбы с зашумлениями

Классические методы восстановления изображения. Классические методы цифровой обработки изображений широко применяются для снижения уровня шума и улучшения качества визуальной информации [39]. В их основе лежат локальные преобразования, обеспечивающие сглаживание высокочастотных искажений, устранение случайных изменений яркости и выравнивание однородных областей изображения. К числу таких методов относятся линейные и нелинейные фильтры пространственной области, включая гауссово сглаживание, медианную фильтрацию и морфологические преобразования. Морфологические методы, основанные на операциях эрозии, дилатации и их комбинациях, применяются для удаления мелких артефактов, заполнения разрывов и сглаживания структуры изображения. Они позволяют корректировать форму объектов и подавлять локальные помехи, однако их эффективность существенно зависит от выбора структурирующих элементов и параметров обработки, что ограничивает универсальность применения в условиях изменяющейся сцены.

Применимость классических методов восстановления изображений к анализу дорожных сцен в неблагоприятных погодных условиях ограничена, поскольку физика шумовых искажений носит неоднородный и динамический характер. Визуальные артефакты на изображении приводят к потере мелких

деталей и снижению информативной ценности изображения, что затрудняет последующее обнаружение объектов. В таких условиях методы локальной фильтрации, основанные на фиксированных преобразованиях, часто приводят к чрезмерному сглаживанию изображения и потере важных деталей, в том числе границ транспортных средств и элементов дорожной инфраструктуры. Это обуславливает необходимость использования подходов, способных учитывать нелинейный характер шума и сохранять структурную целостность изображения [40, 41].

Нейросетевые методы восстановления изображения. В отличие от классических подходов, использующих фиксированные преобразования, методы восстановления на основе нейронных сетей обладают обобщающей способностью и формируют адаптивное отображение между зашумлённым и восстановленным изображением на основе обучения на данных [42, 43].

Развитие методов восстановления изображений обусловлено не только необходимостью повышения их визуального качества, но и задачей обеспечения устойчивости алгоритмов компьютерного зрения в неблагоприятных условиях наблюдения. В [44] восстановление используется как этап предварительной обработки, направленный на повышение точности обнаружения объектов и устойчивости последующего анализа дорожной сцены. В исследованиях [45–47] предложены методы, объединяющие восстановление и распознавание изображений, что позволяет учитывать воздействие искажений непосредственно при формировании информативных признаков.

Среди нейросетевых методов восстановления можно выделить подходы на основе сверточных нейронных сетей, ориентированные преимущественно на восстановление локальных текстур и границ объектов, генеративные модели, учитывающие структуру сцены, а также методы с итеративным процессом реконструкции изображения. Различия в принципах обработки определяют особенности применения данных методов к отдельным видам атмосферных искажений [48].

Методы восстановления на основе свертки являются базовым инструментом восстановления изображений благодаря сравнительно высокой вычислительной эффективности и способности работать с потоковыми видеоданными. Основное преимущество таких методов заключается в извлечении локальных признаков изображения, позволяющих эффективно подавлять шумовые структуры, характерные для атмосферных осадков [49, 50].

Итеративная обработка представляет собой один из подходов к построению нейросетевых методов восстановления изображений [51]. Вместо однократного преобразования исходного изображения применяется рекуррентная схема, в рамках которой промежуточный результат последовательно уточняется на протяжении нескольких итераций. Повторное использование вычислительных блоков позволяет уменьшить требуемую глубину сети, при этом удаление дождевых артефактов осуществляется без существенного увеличения вычислительных затрат. В [52] последовательная обработка организована иным образом: процесс восстановления разделен на взаимосвязанные этапы, работающие с признаками различных масштабов, что позволяет одновременно учитывать общий контекст изображения и локальные текстурные характеристики при наличии пространственно неоднородных атмосферных искажений.

Разделение отдельных операций восстановления реализовано в [53]. Структуры, соответствующие дождевым артефактам, сначала выделяются из изображения, после чего выполняется реконструкция деталей сцены. Предварительное подавление артефактов снижает их влияние на последующую обработку и позволяет ограничить сглаживание границ объектов.

Для изображений, полученных в условиях тумана, применяются модели, учитывающие особенности атмосферного рассеяния. В [54] оценка параметров атмосферы непосредственно включена в процедуру восстановления, что позволяет сократить количество отдельных этапов обработки. Относительно низкая вычислительная сложность метода делает его пригодным для использования в системах реального времени. Вместе с тем эффективность

удаления тумана зависит от того, насколько условия, представленные в обучающих данных, соответствуют характеристикам обрабатываемой сцены [55].

Несмотря на широкое распространение, сверточные методы преимущественно ориентированы на локальные особенности изображения и испытывают трудности при обработке деградаций сцены.

Методы восстановления на основе свертки и блоков внимания являются следствием развития сверточных нейросетевых архитектур для адаптивной обработки различных типов деградации. Такие подходы стремятся компенсировать ограниченность локального рецептивного поля CNN без существенного увеличения вычислительной сложности [56].

Сочетание механизмов внимания, функционирующих на различных уровнях, используется в [57]. Предложенный метод способен обрабатывать искажения, различающиеся по пространственной структуре и масштабу посредством применения локальных внимания и глобальных механизмов внимания.

В [58] предложен метод учета неоднородного характера атмосферных воздействий посредством извлечения признаков параллельно в разных цветовых пространствах, что позволяет отдельно учитывать характеристики цвета, яркости и насыщенности, которые могут изменяться в условиях плотного тумана и неравномерного освещения.

В [59] предложен метод динамически формируемых фильтров на основе свертки и блоков внимания, адаптирующихся к содержимому изображения. Важной особенностью такого метода является возможность учитывать различия между реальными и синтетическими изображениями. Параметры этих фильтров определяются характеристиками обрабатываемой сцены.

Масштаб признаков также учитывается в архитектуре, предложенной в [60], где выходы сверточных операций обработки локальных текстурных признаков объединяются с глобальными зависимостями блоков внимания. В

результате восстановление деталей изображения выполняется с учетом общего содержания сцены при сохранении умеренных вычислительных затрат.

Ограничением данной группы методов является обработка локальных признаков, ограничивающая учет глобальных взаимосвязей объектов на изображении.

Методы восстановления на основе генеративно-состязательных сетей получили широкое распространение в задачах восстановления изображений благодаря способности этих сетей формировать визуально правдоподобные структуры сцены. Концепция обучения была предложена авторами работы [61]. В отличие от сверточных нейронных сетей, ориентированных преимущественно на минимизацию попиксельного шума, генеративно-состязательные сети (GAN) обучаются воспроизводить распределение «чистых» изображений, что позволяет сохранить текстурные детали объектов и уменьшить эффект избыточного сглаживания, часто возникающий при устранении погодных артефактов.

Одним из первых подходов к применению GAN для удаления осадков стал метод, представленный в работе [62]. Согласно авторам работы восстановление изображения осуществляется генератором, использующим различные типы связей, тогда как дискриминатор анализирует характеристики изображения на разных масштабах. Схема состязательного обучения призвана обуславливать нейронную сеть формировать визуально реалистичные результаты.

В GAN стали включаться механизмы внимания, предназначенные для определения областей изображения, наиболее подверженных искажениям. В [63] карта внимания последовательно уточняется с использованием рекуррентных и сверточных слоев. Выделенные области затем обрабатываются контекстным автокодировщиком, что позволяет модели сосредоточить основное воздействие на участках, содержащих дождевые артефакты, сохраняя при этом информацию в остальных областях изображения.

В [64] требование к использованию парных обучающих данных устраняется за счет применения циклической согласованности. В этом случае обучение может проводиться на несопоставленных изображениях из различных

доменов. Дополнительное выделение структур, связанных с дождем, совместно с перцептивной функцией потерь используется для сохранения мелких деталей и облегчения переноса модели на реальные сцены.

Более общий подход к доменному преобразованию рассматривается в [65], где используется единая схема обработки различных погодных зашумлений посредством применения выравнивания признаков и построения каскадных блоков внимания.

Наряду с повышением качества восстановления исследования также направлены на оценку качества задач последующего обнаружения объектов на восстанавливаемом изображении. В [44] метод на основе U-Net в процессе обучения учитывает высокоуровневые признаки восстанавливаемого изображения за счет дополнения механизмом самовнимания.

Методы на основе GAN обеспечивают высокое качество восстановления изображений при различных типах атмосферных искажений и направлены на сохранение структурных особенностей изображения при сопоставимо высокой скорости обработки.

Методы восстановления на основе трансформеров позволяют учитывать признаковые связи между удалёнными областями изображения за счет механизма внимания.

Для снижения вычислительной сложности самовнимания применительно ко всему изображению в [66] применяется сочетание канального внимания и иерархической архитектуры «кодировщик-декодировщик». Операции свертки обеспечивают выделение локальных характеристик изображения, а объединение признаков между уровнями сети позволяет учитывать более широкий контекст сцены.

Авторы [67] представили метод, включающий предварительное обучение нейронной сети на обширном наборе данных, состоящем из пар синтетически искаженных изображений, с последующей адаптацией к конкретной предметной области в рамках задачи восстановления. Дальнейшее обобщение этого подхода

предложено в [68]. Для определенных типов искажений определяются специализированные веса, которые дополняются низкоранговыми сдвигами.

Иерархическая обработка также применяется в [69]. Кодировщик и декодировщик связаны на соответствующих уровнях, а самовнимание вычисляется в пределах неперекрывающихся окон. Для карт признаков высокого разрешения такой подход позволяет ограничить вычислительные затраты, тогда как после уменьшения пространственных размеров признаков область, соответствующая одному окну, охватывает более значительную часть исходной сцены. При этом потеря деталей при переходах между масштабами компенсируется обучаемым модулем в декодировщике.

Для удаления снежных осадков в [70] сначала формируется приближенная маска искажений с использованием вариационного автоэнкодера. На основе этой маски задается ограниченный набор запросов внимания, благодаря чему обработка концентрируется на поврежденных областях изображения. Подобный способ организации запросов предложен в [71] – запросы формируются на основе многомасштабных признаков и сопоставляются с локальными областями изображения посредством перекрестного внимания.

Сочетание сверточных и трансформерных блоков представлено в [72]. Сверточные компоненты кодировщика и декодировщика используются для извлечения и восстановления локальных признаков, тогда как трансформер работает при уменьшенном пространственном разрешении.

В [73] предложен подход к построению универсальных моделей на явном задании информации о типе искажения. Для этого используется модуль генерации промптов, который выделяет признаки, соответствующие определенному типу искажения, и передает их в трансформер восстановления. Подобное решение предложено авторами [74], где процесс восстановления управляется с помощью текстовой инструкции, а ее представление определяет операцию, выполняемую единой моделью при удалении дождя, шума, размытия или тумана.

Трансформерные методы различаются способом ограничения внимания. Эти приёмы снижают затраты по сравнению с прямым вниманием на полном изображении, но выбор метода для видеопотока требует отдельно учитывать время обработки кадра и переносимость на реальные погодные условия.

Методы восстановления на основе диффузии представляют собой класс современных генеративных подходов, в которых реконструкция изображения выполняется путем последовательного обращения процесса его искажения посредством итеративного удаления шума [75, 76]. В отличие от GAN, предполагающей совместное обучение генератора и дискриминатора, методы на основе диффузии описывают вероятностный переход от деградированных данных к восстановленному изображению [77]. Последовательный характер процесса реконструкции способствует стабильности обучения и сохранению текстурной информации при наличии сложных атмосферных искажений.

Универсальный метод диффузионной очистки, представленный в [78], представляет процесс обратной диффузии с определением параметров восстановления с учетом интенсивности искажений, а совокупность применения многомасштабной оценки зашумления позволяет обрабатывать артефакты осадков различной выраженности. Авторы [79] развивают идею методов диффузионного восстановления от комбинированных атмосферных осадков. Семантические запросы, формируемые на основе характеристик деградации и содержимого сцены, и интеграция признаков через механизм кросс-внимания обеспечивает адаптацию метода к артефактам.

В работе [80] предложен патчевый диффузионный метод, использующий разбиение изображения на перекрывающиеся фрагменты, каждый из которых обрабатывается независимо с последующим объединением результатов. Полученный результат позволяет многократно уточнять локальные области сцены и сохранять мелкие детали изображения при различных видах погодных зашумлений, при этом теряя в скорости обработки и требуя больших ресурсов.

В работе [80] предложен метод диффузии на основе патчей, который предполагает разделение исходного изображения на перекрывающиеся участки,

их независимую обработку и последующее объединение. Анализ патчей обеспечивает постепенное уточнение локальных областей и позволяет сохранить мелкие детали в условиях различных погодных искажений, однако ограничением такого метода является вычислительная сложность многократной обработки отдельных фрагментов и время восстановления изображения.

Диффузионные методы позволяют восстанавливать изображения при сложных атмосферных искажениях и сохранять детали изображения. Вместе с тем восстановление выполняется в ходе последовательности итераций, каждая из которых требует дополнительного вычислительного шага. В результате время обработки и потребление вычислительных ресурсов возрастают, что затрудняет применение таких моделей при непрерывной обработке видеопотока [81–83]

1.6. Выводы по главе

В главе выполнен анализ современных методов автоматического обнаружения дорожно-транспортных происшествий на основе обработки визуальной информации, включающий исследование подходов к анализу дорожных сцен, влияния неблагоприятных погодных условий на качество распознавания и современных методов восстановления изображений.

Установлено, что видеопотоки стационарных камер являются наиболее приемлемым источником информации для масштабируемого автоматического обнаружения дорожно-транспортных происшествий в городской среде.

Выявлено, что методы анализа отдельных кадров и методы пространственно-временного анализа решают взаимодополняющие задачи и характеризуются различными источниками ошибок. Следовательно, достоверное обнаружение ДТП требует исследования возможностей использования пространственных признаков изображения и динамических характеристик дорожной сцены.

Показано, что неблагоприятные погодные зашумления снижают контрастность изображения, искажают границы объектов, формируют ложные текстурные признаки и частично скрывают элементы сцены, что отрицательно влияет на обнаружение объектов и событий на изображении.

Установлено, что обработка видеопотоков накладывает ограничения как на качество восстановления, так и на вычислительную сложность используемых методов. Существенное значение имеет сохранение признаков, используемых на последующих этапах детектирования, классификации и анализа взаимного положения транспортных средств. Среди рассмотренных подходов генеративные методы сочетают в себе способность восстанавливать структуру изображения с меньшими затратами времени на обработку по сравнению с итерационными методами, что делает их перспективными для дальнейших исследований в области потоковой видеоаналитики.

Проведенный анализ позволил сформулировать требования к разрабатываемым методам обработки визуальной информации и обосновать направления дальнейшего исследования. Полученные результаты обосновывают необходимость дальнейших глав диссертационного исследования.

Глава 2. Исследование и разработка метода определения дорожно-транспортных происшествий.

Во второй главе диссертационного исследования решается задача исследования и разработки метода обнаружения дорожно-транспортных происшествий по данным видеокамер. В соответствии с результатами первой главы разрабатываемый метод должен обеспечивать выявление информативных признаков аварийной ситуации при обработке множественных видеопотоков и в режиме реального времени.

На первом этапе рассматривается задача классификации отдельных кадров видеопотока, позволяющая определять наличие визуальных признаков дорожно-транспортного происшествия без анализа траекторий движения объектов. Для выбора базового метода классификации ДТП выполняется сравнительный анализ нейросетевых классификаторов (Mobile Net, Efficient Net, DCNN, ConvNeXtv2), различающихся параметрической емкостью, вычислительной сложностью и способом извлечения пространственных признаков изображения.

На втором этапе исследуются особенности выбранного нейросетевого классификатора и разрабатывается нейросетевой классификатор изображений, ориентированный на извлечение информативных признаков дорожной сцены при сохранении приемлемой вычислительной эффективности.

На третьем этапе разрабатывается метод обнаружения ДТП на видеопотоке с двухэтапным принятием решения об обнаружении происшествия, включающим формирование событий-кандидатов по результатам анализа динамики транспортных средств и их последующую нейросетевую классификацию.

2.1. Исследование методов определения ДТП

Автоматическое определение дорожно-транспортных происшествий по видеопотоку предполагает выявление изменений в дорожной обстановке, соответствующих возникновению аварийной ситуации, признаки которой могут проявляться как в пространственной структуре отдельных изображений, так и в характере изменения положения транспортных средств с течением времени.

Как показано в первой главе, каждый из рассмотренных подходов имеет свои ограничения. Классификация отдельных кадров меньше зависит от ошибок сопровождения объектов, однако не учитывает динамику события. Анализ последовательности кадров позволяет учитывать развитие дорожной ситуации, но зависит от устойчивости детектирования объектов.

Поэтому исследование начинается с классификации отдельных изображений. Это позволяет оценить возможность распознавания визуальных признаков ДТП, выбрать базовый классификатор и определить способ улучшения извлечения пространственных признаков.

Несмотря на высокие показатели качества, достигаемые современными методами на основе механизма самовнимания [84, 85, 86], их применение сопровождается существенным увеличением скорости обработки данных и объема требуемых вычислительных ресурсов. В связи с этим дальнейшее исследование ограничено сверточными классификаторами, окупающих качество распознавания визуальных признаков ДТП с невысокой вычислительной сложностью и возможностью обработки кадров в составе метода анализа видеопотока. Далее в главе рассматриваются классификаторы MobileNet, EfficientNet разных конфигураций, DCNN и семейство ConvNeXtV2, представляющие различные принципы построения сверточных классификаторов и отличающиеся параметрической емкостью, структурой вычислительных блоков и способом масштабирования модели.

Сбор и подготовка данных. Для оценки выбранных классификаторов было необходимо подготовить данные, содержащие сцены дорожных инцидентов. Для этого были проанализированы и отобраны несколько источников данных, включающих как открытые готовые наборы данных, так и самостоятельно собранные данные из открытых источников. Изображения содержат как синтетические, так и реальные сцены ДТП. Целью подготовки данных было формирование набора данных, отражающей условия городского трафика.

Собранный набор состоит из данных, представленных в нескольких источниках. Первая часть данных взята из Traffic-Net [87], включающего в себя реальные изображения четырех классов, каждый из которых содержит 1100 изображений: ДТП, пожар, плотный трафик и разреженный трафик. Данные, содержащие сцены пожара и плотного трафика были отброшены поскольку не релевантны в рамках решаемой задачи, оставлены сцены, содержащие ДТП и разреженный трафик.

Вторая часть представлена в [88] и содержит несколько поднаборов синтетических видеороликов, на которых смоделированы различные сценарии, среди которых отобран поднабор «crash-likelihood». Данная выборка также содержит фрагменты как с происшествиями, так и без них. Из данного набора данных отобрано более 5000 изображений для обучения, более 1000 для валидации и 2500 для тестирования выбранных методов. Синтетические данные призваны расширить итоговый набор данных поскольку реальных изображений среди открытых источников содержится мало, как с камер видеонаблюдения, так и с других ракурсов съемки, а также с целью разбавить различной вариативностью смоделированных сцен.

Третья часть содержит изображения, собранные самостоятельно из открытых источников. Были проанализированы записи с камер дорожного наблюдения, в числе которых содержатся сцены зафиксированных дорожных происшествий и движение без происшествий. По результатам анализа собрано

1026 реальных изображений, содержащих 426 изображений ДТП и 600 изображений безаварийного дорожного трафика.

Итоговое количество данных составило 12426 изображений. Выборки разделялись на обучающую, валидационную и тестовую части в соотношении 60/15/25. Большое количество данных в тестовой выборке способствует подтверждению валидности оценки при различных условиях получения изображений.

Исследование классификаторов. Все исследуемые нейросетевые классификаторы были дообучены на собранном наборе данных. Использовался единый подход без заморозки слоев, предобученные веса обучались 30 эпох, скорость обучения $5 \cdot 10^{-5}$, оптимизатор AdamW, размер изображения 224x224, функция потерь CrossEntropyLoss, размер батча 8.

Полученные результаты сравнивались по следующим показателям: Accuracy, Precision, Recall, F1, время обработки кадра на графическом ускорителе. Исследование производительности проводилось на сервере с процессором AMD EPYC 7742, 64 ядра, 128 потоков и видеокартой NVIDIA Tesla A100-SXM4-40GB.

Таблица 2.1.1 Результаты сравнения методов определения ДТП

Метод	Accuracy	Precision	Recall	F1	Время обработки, мс
MobileNet [17]	0.91	0.89	0.94	0.92	7
EfficientNetB1 [16]	0.9	0.94	0.92	0.9	10
EfficientNetB7 [16]	0.87	0.91	0.84	0.87	30
DCNN [18]	0.86	0.85	0.88	0.86	10
ConvNeXtV2 Atto	0.86	0.87	0.84	0.83	6
ConvNeXtV2 Femto	0.87	0.92	0.88	0.87	6
ConvNeXtV2 Pico	0.86	0.86	0.89	0.83	6
ConvNeXtV2 Nano	0.87	0.89	0.84	0.83	10
ConvNeXtV2 Tiny	0.84	0.82	0.90	0.85	10

На рисунке 2.1 представлены кривые изменения метрик качества классификации изображений исследуемых методов на обучающей и валидационной выборках.

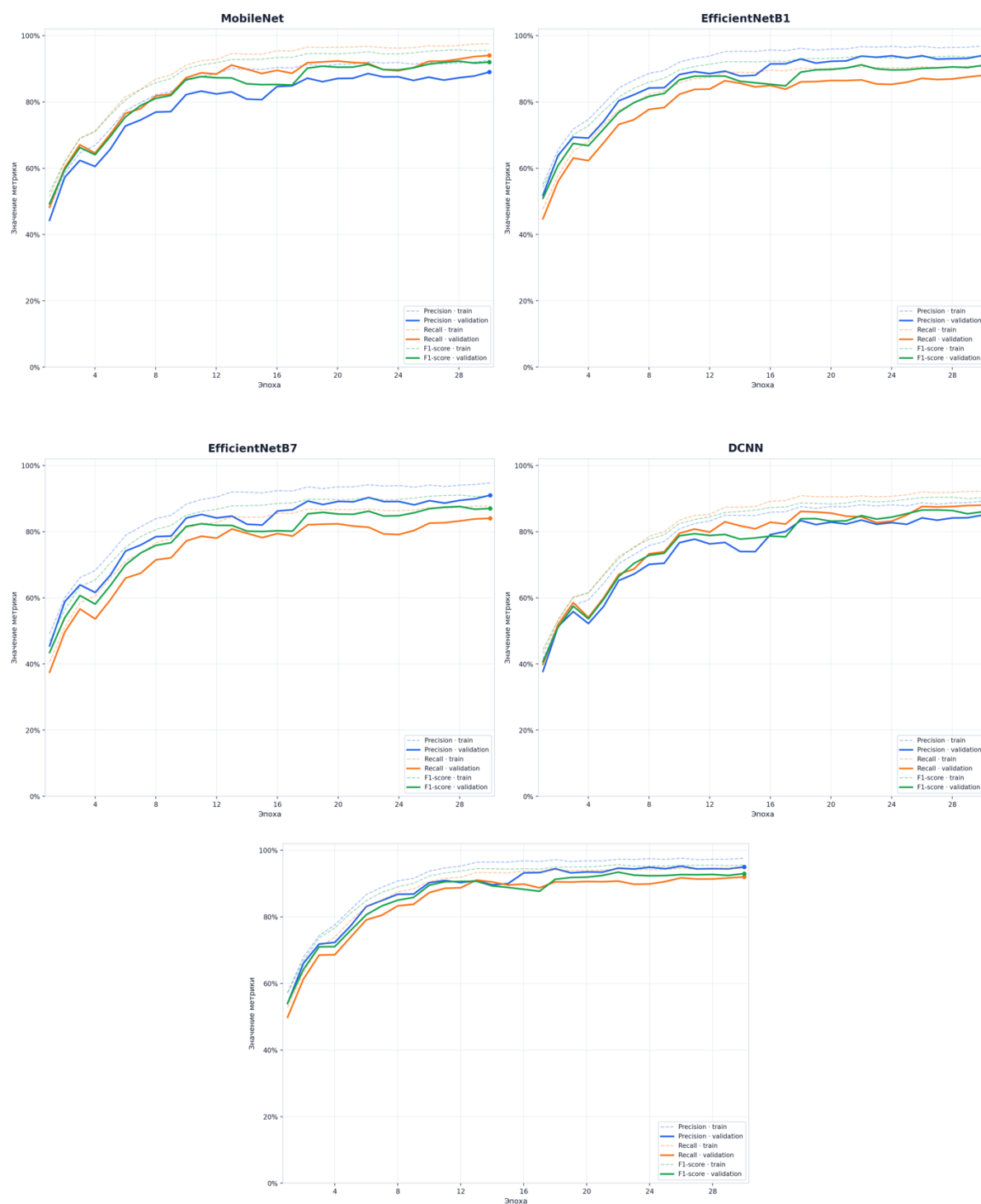


Рисунок 2.1 – Графики со значениями метрик точности на обучающей и валидационной выборках исследуемых классификаторов

Более компактные классификаторы демонстрируют высокую скорость обработки, однако уступают по полноте и устойчивости классификации более ёмким нейронным сетям. Поскольку в задаче обнаружения ДТП требуется как быстрая обработка кадров, так и высокая точность классификации, сверточные нейронные сети остаются наиболее подходящим выбором для таких задач благодаря низкой задержке обработки изображения [89, 90, 91]. При выборе классификатора также необходимо учитывать качество распознавания, поскольку пропуски ДТП и ложные срабатывания влияют на результат автоматического мониторинга. Хотя методы компьютерного зрения на основе трансформеров, получившие широкое распространение в задачах компьютерного зрения с конца прошлого десятилетия в научном сообществе, продемонстрировали высокую точность за счет глобального взаимодействия признаков, тем не менее обычно требуют значительно более высоких вычислительных затрат, которые ограничивают их применимость в реальном времени. ConvNeXt [92] сочетает ряд принципов, характерных для визуальных трансформеров, с полностью сверточной структурой. Использование увеличенных ядер свертки, нормализации слоев и измененной схемы агрегации признаков позволяет учитывать более широкий пространственный контекст, что дает возможность применения ConvNeXtV2 в системах обнаружения происшествий. Авторы оригинальной статьи [92] представили обширные экспериментальные результаты. Семейство ConvNeXt было обучено и протестировано на наборах данных ImageNet-22K и ImageNet-1K, и результаты сопоставимы с результатами трансформеров, часто достигая более высокой точности с лучшей пропускной способностью. ConvNeXt также была протестирована на задачах обнаружения и сегментации объектов с использованием набора данных COCO. Она была предварительно обучена с использованием Mask R-CNN, продемонстрировав высокую производительность, сопоставимую с современными трансформерами, которые значительно превосходят ConvNeXt по количеству параметров.

Одним из основных нововведений в ConvNeXtV2 является использование полностью сверточного маскированного энкодера (FCMAE) для самообучения. В отличие от оригинального маскированного энкодера (MAE), в котором использовался асимметричный кодировщик-декодировщик на основе механизма внимания [93], FCMAE полностью основан на сверточных операциях. Маскированный вход обрабатывается как набор разреженных патчей с использованием разреженной свертки, что позволяет обрабатывать только видимые части изображения, повышая эффективность обучения. Декодировщик заменен одним сверточным блоком ConvNeXt, что делает всю нейронную сеть полностью сверточной.

На основе анализа таблицы 2.1.1 конфигурация Femto была взята за основу как базовая конфигурация для разработки классификатора по соотношению меньшего количества параметров и лучших показателей в своем семействе.

2.2. Разработка метода обнаружения ДТП

Ограниченность локального характера стандартных сверточных операций может приводить к потере информации о глобальных взаимосвязях между элементами сцены [94, 95], имеющих значение при выявлении аварийных ситуаций. В связи с этим разработка метода направлена на повышение информативности признаков, используемых для классификации состояния дорожной сцены. В рамках разработки предложен способ формирования признакового представления изображения, основанный на совместном анализе пространственных и спектральных характеристик визуальной информации. Для реализации данного подхода использованы операции быстрого преобразования Фурье (FFC) [96], позволяющие учитывать глобальные компоненты структуры изображения при обработке высокоуровневых признаков.

Характерные визуальные признаки ДТП формируются не только в локальных областях изображения, но и на уровне всей сцены. Помимо кузовных повреждений кадр содержит нетипичное расположение транспортных средств, нарушение общей геометрии движения. В таких условиях применение классической операции свертки, имеющей ограниченное рецептивное поле, также ограничивает и выявление таких глобальных и полу-глобальных зависимостей. В свою очередь операция свертки на основе FFC позволяет обрабатывать изображение с расширением рецептивного поля и объединением информации из разных масштабов, при этом фиксировать глобальный контекст в рамках одной операции, не увеличивая глубину сети и вычислительную сложность. В отличие от задач объектного распознавания, признаки дорожно-транспортных происшествий проявляются не только в локальных текстурах, но и в нарушении глобальной структуры сцены и взаимного расположения объектов.

Авторы [96] заявляют, что ключевым компонентом FFC является спектральный преобразователь, который направлен на эффективное увеличение рецептивного поля операции свертки для полного разрешения тензора входных характеристик с помощью дискретного преобразования Фурье. Сначала входной тензор, представленный как $X \in R^{H \times W \times C}$, где $H \times W, C$ – пространственное разрешение и количество каналов, разделяется на пространственную часть $X^l \in R^{H \times W \times (1-\alpha_{in})C}$ и спектральную часть $X^g \in R^{H \times W \times \alpha_{in}C}$. $\alpha_{in} \in [0, 1]$ представляет процент каналов признаков и контролирует вклад локальной и глобальной информации на разных уровнях сети. Выходной вектор $Y \in R^{H \times W \times C}$ аналогичным образом разлагается на локальную и глобальную части, и соотношение глобальной части выходного тензора к глобальной части также определяется гиперпараметром $\alpha_{out} \in [0, 1]$. Таким образом, операция быстрой свертки Фурье описывается как

$$Y^l = Y^{l \rightarrow l} + Y^{g \rightarrow l} = f_l(X^l) + f_{g \rightarrow l}(X^g), \quad (2.1)$$

$$Y^g = Y^{g \rightarrow g} + Y^{l \rightarrow g} = f_g(X^g) + f_{l \rightarrow g}(X^l), \quad (2.2)$$

где $Y^{l \rightarrow l}, Y^{l \rightarrow g}, Y^{g \rightarrow l}$ применяют обычную свертку, а $Y^{g \rightarrow g}$ применяет дискретное преобразование Фурье.

С точки зрения сложности, FFC сопоставим с обычными операциями свертки по количеству параметров и операций. Важное преимущество FFC проявляется при использовании больших ядер, поскольку спектральный преобразователь может быть обучен с помощью глобального рецептивного поля с использованием сверточного ядра 1×1 .

В ConvNeXt одним из ключевых элементов сети является использование увеличенных ядер операций свертки размером 7×7 , встроенной в ограничивающие блоки. Хотя авторы ConvNeXt показали, что расширение ядра свёртки позволяет существенно увеличивать эффективное рецептивное поле, улучшая способность учитывать более широкий контекст сцены без роста вычислительной сложности, ядра такого размера являются инструментом локального анализа. Они эффективно охватывают ближайшие окрестности пикселя, будучи ограниченными фиксированным пространственным окном. Характерные признаки ДТП при анализе видеопотока имеют преимущественно глобальную или полу-глобальную природу, выраженную изменением взаимного расположения транспортных средств на больших расстояниях, нарушением траекторий и разрывом потоков движения. Для таких случаев требуется механизм, охватывающий всю сцену и обобщающий зависимости между удаленными областями изображения.

Разработанный нейросетевой классификатор представляет способ формирования признакового представления дорожной сцены, предусматривающий использование блока Fast Fourier Convolution в третьей стадии архитектуры ConvNeXtV2 вместо пространственной свёртки с размером ядра 7×7 . Первые две стадии сети выполняют извлечение низко- и среднеуровневых признаков изображения без изменений. На третьей стадии обработка карт признаков осуществляется блоком FFC, в котором вычисления выполняются как в пространственной, так и в частотной областях. Спектральная обработка реализуется посредством прямого преобразования Фурье карт

признаков, выполнения операций свёртки над спектральным представлением и последующего обратного преобразования в пространственную область. Полученные признаки используются последующими вычислительными блоками сети и классификационной головой для прогнозирования наличия дорожно-транспортного происшествия.

Разработанный классификатор был обучен на собранном наборе данных с идентичным предыдущему эксперименту набором гиперпараметров. В таблице 2.2.1 представлены результаты тестирования модифицированной и базовых методов определения ДТП.

Таблица 2.2.1. Результаты сравнения методов определения ДТП

	Accuracy	Precision	Recall	F1-мера	Время обработки, мс
MobileNet [17]	0.91	0.89	0.94	0.92	10
EfficientNet1 [16]	0.9	0.94	0.88	0.91	10
EfficientNet7 [16]	0.87	0.91	0.84	0.87	30
DCNN [18]	0.86	0.85	0.88	0.86	10
CoNvNeXtV2	0.94	0.95	0.92	0.93	6
Разработанный классификатор	0.94	0.94	0.96	0.95	4

На рисунке 2.2 представлены кривые изменения метрик качества исследуемого и разработанного классификаторов в процессе обучения.

На рисунке 2.3 показана матрица ошибок, полученная после обучения разработанного классификатора.

По результатам, представленным в таблице 2.2.1 видно, что разработанный классификатор превосходит сравниваемые классификаторы включая базовый как по метрикам точности обнаружения ДТП, так и по скорости обработки одного изображения.

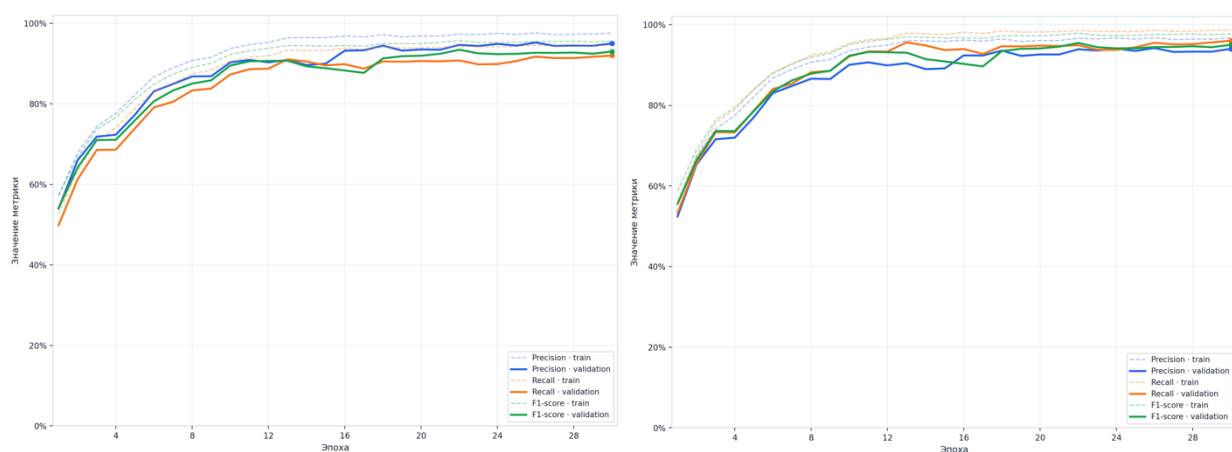


Рисунок 2.2 - Графики со значениями метрик Precision, Recall, F1 на обучающей и валидационной выборках классификаторов ConvNeXtV2 (слева) и разработанного классификатора (справа)

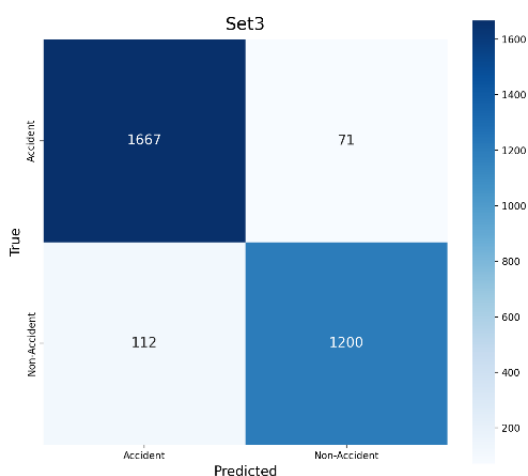


Рисунок 2.3 – Матрица ошибок разработанного метода определения ДТП

Однако для корректного определения происшествий в условиях обработки видеопотока опираться только на классификацию отдельных кадров может быть недостаточно, поскольку признаки аварийности появляются в динамике движения объектов и формируются во времени. Для этого в данной работе предложен метод обнаружения ДТП на основе двухстадийного подхода с анализом динамики сцены: на первой стадии происходит локализация транспортных средств и оценка их параметров движения – скорости, ускорения, траектории – с целью фильтрации ложных и выявления потенциальных событий-

кандидатов, на второй – при достижении порогового значения применяется разработанный классификатор для получения вероятностного прогноза о наличии ДТП на изображении по предложенным кандидатам.

Локализация участников дорожного движения выполняется с помощью алгоритма YOLOv8m, так как он является современным нейросетевым детектором, применяемым в системах реального времени за счет проведения локализации и классификации объектов за один проход по изображению.

Для ассоциации участников движения используется локальная схема сопоставления объектов между соседними кадрами, основанная на минимизации евклидова расстояния между центроидами ограничивающих прямоугольников. Это не обеспечивает устойчивой идентификации объектов на длительном расстоянии в сравнении с алгоритмами трекинга, в том числе позволяющих восстанавливать идентификацию автомобилей после частичных перекрытий [97, 98], но достаточно для оценки краткосрочных изменений движения и вычислительно дешевле.

В основе метода обнаружения ДТП лежит анализ движения транспортных средств по последовательности кадров. Аварийное событие проявляется через изменения скорости, ускорения, направления движения и взаимного положения автомобилей. Наиболее характерны резкое торможение при попутном движении, боковые столкновения на перекрестках, столкновения при перестроении, а также ситуации, предшествующие ДТП и сопровождающиеся резким изменением кинематических параметров. Эти признаки соотносятся со сценариями развития аварий, приведенными в модели угроз первой главы диссертационного исследования.

Поскольку наблюдаемые кинематические характеристики имеют различную природу необходимо учитывать влияние на общую оценку наличия аномального события. Для этого предложено ввести весовые коэффициенты, выполняющие роль нормирующих и управляющих параметров и позволяющие адаптировать метод под разные условия входных данных.

Шаг 1. Локализация транспортных средств. На каждом кадре I_t выполняется локализация транспортных средств с использованием алгоритма YOLO. На выходе образуется множество ограничительных рамок:

$$B^t = \{b_1^t, b_2^t, \dots, b_{N_t}^t\}, \quad (2.3)$$

где N_t – число обнаруженных транспортных средств на кадре t , b_i^t – ограничивающая рамка транспортных средств.

Шаг 2. Вычисление центроид транспортных средств. Для каждого транспортного средства определяется его центроид:

$$c_i^t = (x_i^t, y_i^t), \quad (2.4)$$

где x_i^t, y_i^t – координаты центроида объекта i .

Центроиды используются в качестве точек отслеживания характеристик движения.

Шаг 3. Отслеживание траекторий и вычисление кинематических параметров. Для каждого транспортного средства на кадре t выполняется сопоставление центроид между соседними кадрами. Соответствие определяется как минимизация евклидова расстояния между центроидами:

$$j = \arg \min_k \|c_i^t - c_i^{t+1}\|, \text{ при } \|c_i^t - c_j^{t+1}\| < d_{max}, \quad (2.5)$$

где d_{max} – максимально допустимое расстояние ассоциации.

Дискретная оценка скорости объекта определяется как вектор:

$$\vec{v}_i^t = \frac{c_j^t - c_i^{t-1}}{\Delta t}. \quad (2.6)$$

Дискретная оценка ускорения определяется как:

$$\vec{a}_i^t = \frac{\vec{v}_i^t - \vec{v}_i^{t-1}}{\Delta t}. \quad (2.7)$$

Шаг 4. Нормирование и сглаживание. Производится нормирование для анализа изменений движения без привязки к направлению:

$$\hat{v}_i^t = \frac{\|\vec{v}_i^t\|}{T_v}, \quad \hat{a}_i^t = \frac{\|\vec{a}_i^t\|}{T_a}, \quad (2.8)$$

где T_v и T_a – пороговые значения для скорости и ускорения соответственно.

Для фильтрации шумов используется экспоненциальное сглаживание:

$$\tilde{x}_i^t = \alpha x_i^t + (1 - \alpha) \tilde{x}_i^{t-1}, 0 < \alpha < 1. \quad (2.9)$$

Шаг 5. Оценка вероятности коллизии по кинематическому анализу.

Для выявления аномальной динамики движения вводятся признаки, отражающие изменения состояния объектов. Падение скорости определяется как:

$$\Delta v_i^t = \max(0, \|\vec{v}_i^{t-1}\| - \|\vec{v}_i^t\|). \quad (2.10)$$

Изменение направления движения определяется как угол между векторами скорости:

$$\Delta \theta_i^t = \arccos\left(\frac{\vec{v}_i^t \cdot \vec{v}_i^{t+1}}{\|\vec{v}_i^t\| \cdot \|\vec{v}_i^{t+1}\|}\right). \quad (2.11)$$

На основе этих признаков вычисляется коэффициент потенциальной аварийной ситуации:

$$K_i^t = \omega_1 \frac{\Delta v_i^t}{T_v} + \omega_2 \hat{a}_i^t + \omega_3 \frac{\Delta \theta_i^t}{T_\theta}, \quad (2.12)$$

где ω_1 , ω_2 , ω_3 – весовые коэффициенты, T_θ – порог изменения направления движения.

Если коэффициент $K_t^{(i)}$ достигает порогового значения T_{pred} , то событие рассматривается как кандидат на наличие ДТП.

Шаг 6. Классификация события разработанным нейросетевым методом определения ДТП. Для проверки наличия или отсутствия предполагаемого события выполняется классификация с применением разработанного метода определения ДТП:

$$p_t = f_{cls}(I^t), \quad (2.13)$$

где $p_t \in [0, 1]$ – вероятность того, что изображение содержит ДТП.

Если коэффициент p_t достигает порогового значения T_{cls} , то изображение считается содержащим признаки происшествия.

Шаг 7. Валидация события по временной последовательности. Для проверки корректности прогноза метода определения ДТП применяется метод подтверждения события в скользящем окне W_t :

$$P_w^t = \sum_{\tau \in W_t} 1(p^\tau > T_{cls}). \quad (2.14)$$

Если коэффициент P_w^t достигает порогового значения T_w , то событие считается обнаруженным.

Алгоритмическое обеспечение метода обнаружения дорожно-транспортных происшествий определяет последовательность выполнения вычислительных операций, переходы между этапами и формирование итогового решения о наличии ДТП. Если описание метода определяет используемые признаки и критерии их оценки, то алгоритмическая реализация отражает последовательность выполнения этих операций в процессе анализа видеоданных. Блок-схема алгоритмического обеспечения разработанного метода представлена на рисунке 2.4.

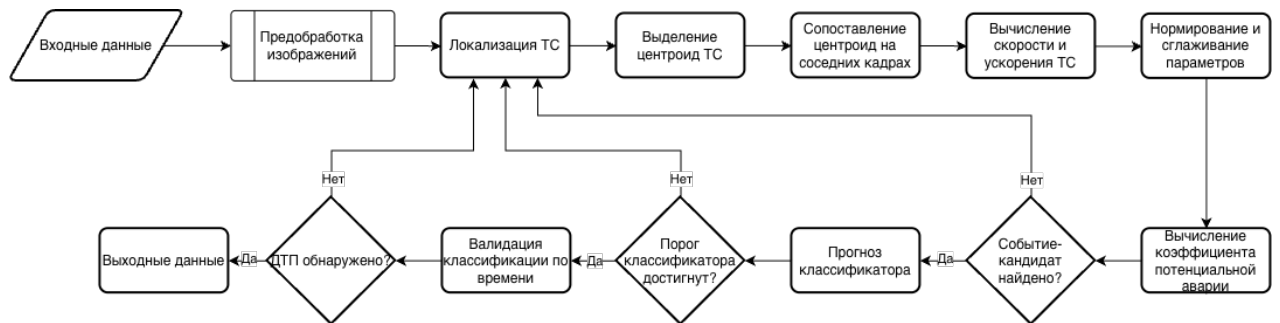


Рисунок 2.4 – Блок-схема алгоритма обнаружения дорожно-транспортных происшествий

Обработка видеопотока выполняется непрерывно, для каждого нового кадра обнаруживаются транспортные средства, после чего анализируются параметры их движения и оценивается текущая дорожная ситуация. Если рассчитанные динамические характеристики не указывают на возможное ДТП, выполняется переход к следующему кадру. При превышении заданного порога запускается нейросетевая классификация, которая проверяет наличие визуальных признаков аварии. Результат классификации дополнительно проверяется на последовательности кадров. Проверка во временном окне

необходима для отделения устойчивого аварийного события от кратковременных изменений дорожной сцены.

Для оценки разработанного метода обнаружения ДТП были проанализированы различные наборы данных, содержащие видеоклипы с нормальным дорожным движением и аномалиями [5, 12, 99]. В них представлены видеофрагменты различной длины, полученные как с видеорегистраторов, так и с дорожных камер видеонаблюдения, полученные в различных условиях перспективы съемки. Так как метод опирается на кинематические параметры, оцениваемые визуально по пикселям, а событие-кандидат определяется с учетом весовых параметров, корректная работа определяется конфигурированием под условия эксплуатации. В связи с этим для проведения оценки работоспособности разработанного метода было принято решение о формировании собственного набора данных, содержащих клипы, приближенные к одинаковым условиям. Сформированный набор включает 50 видеозаписей, содержащих дорожно-транспортные происшествия, и 50 видеозаписей с нормальным дорожным трафиком, длина клипов варьируется от 8 до 15 секунд. Результаты экспериментов представлены в таблице 2.2.2.

Таблица 2.2.2. Результаты экспериментальной оценки метода обнаружения ДТП на видео

Precision	Recall	F1	Среднее время обработки, мс
0.85	0.88	0.86	20

Снижение значений метрик точности по сравнению с результатами отдельного классификатора связано накоплением ошибок на последовательных этапах обработки видеопотока. Результат работы метода зависит не только от классификации изображений, но и от точности обнаружения и сопровождения транспортных средств, а также корректного формирования событий-кандидатов. Полученные значения метрик точности обнаружения ДТП и матрица ошибок на рисунке 2.5 показывают, что разработанный метод обнаруживает большую часть ДТП, сохраняя при этом небольшое число ложных срабатываний.

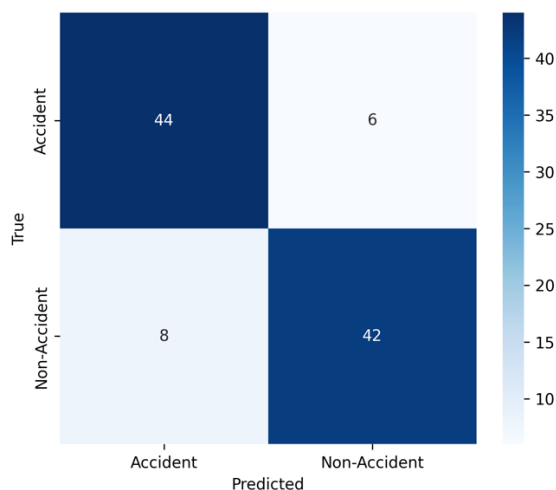


Рисунок 2.5 – Матрица ошибок метода обнаружения ДТП

В результате был разработан метод для обнаружения дорожно-транспортных происшествий, в котором классификация дорожной сцены дополняется анализом движения транспортных средств. Учет изменений кинематических характеристик и взаимного расположения позволяет отличать признаки аварии от кратковременных колебаний транспортного потока. При формировании событий-кандидатов и проверке результатов классификации используются скользящее временное окно, что снижает влияние отдельных ошибок обнаружения и нестабильности видеопотока.

2.3. Выводы по главе

Во второй главе исследованы современные методы обнаружения дорожно-транспортных происшествий по видеоданным. Проведено исследование современных сверточных классификаторов изображений в задаче обнаружения дорожно-транспортных происшествий. По результатам сравнительного анализа обоснован выбор конфигурации для применения в условиях обработки множественных видеопотоков. На ее основе разработан нейросетевой классификатор, предназначенный для обнаружения событий на изображении.

На основе полученного классификатора разработан нейросетевой метод обнаружения дорожно-транспортных происшествий, отличающийся двухэтапным принятием решения об обнаружении происшествия. На первом этапе выполняются локализация транспортных средств на изображении, анализ параметров их движения и формирование событий-кандидатов. На втором этапе выполняется нейросетевая классификация событий-кандидатов на последовательности кадров. Предложенная организация принятия решения позволяет учитывать динамику развития дорожной сцены при анализе видеопоследовательности.

Экспериментальная оценка разработанного метода на независимой тестовой выборке показала значения метрик точности обнаружения ДТП $Precision = 0,85$, $Recall = 0,88$ и $F1 = 0,86$ при времени обработки изображения не более 20 мс. Полученные результаты соответствуют требованиям, сформулированным по итогам первой главы, и являются основой для дальнейших исследований, направленных на повышение устойчивости разработанного метода в условиях зашумлений, вызванных погодными осадками.

Глава 3. Метод восстановления изображения от погодных зашумлений

Третья глава диссертационного исследования посвящена анализу методов восстановления изображений, применяемых в задачах компьютерного зрения для повышения устойчивости алгоритмов к зашумлениям, вызванным атмосферными осадками. Особое внимание уделяется влиянию атмосферных искажений на качество данных и точность последующего обнаружения объектов и событий. Рассматриваются существующие подходы к восстановлению изображений, включая сверточные и генеративные методы, с акцентом на их способности устранять деградации и сохранять структурные особенности сцены. Также проводится анализ исследования методов, включая дистилляцию знаний и квантизацию, на вычислительную эффективность и качество восстановления. Показано, что существующие методы обеспечивают высокое качество визуального восстановления, однако обладают ограничениями, связанными с вычислительной сложностью и недостаточной устойчивостью при обработке видеопотока в реальном времени. На основе проведённых исследований разработан метод восстановления изображений от погодных зашумлений, обеспечивающий сохранение структурных характеристик сцены при сниженной вычислительной сложности и применимость в задачах реального времени.

3.1. Сравнительный анализ методов удаления дождя с изображений

Многочисленные виды атмосферных осадков имеют различную природу, тем самым по-разному оказывая влияние на качество распознавания зашумленных изображений. В данной работе задача повышения качества распознавания образов на изображении решалась путем исследования и разработки метода восстановления изображения от шумов, вызванных дождевыми осадками, поскольку дождь является одним из распространённых

видов атмосферных осадков в умеренном климате. По данным Гидрометцентра России, в Москве среднее число дней в году с осадками более 0,1 мм составляет от 8 до 12 в месяц в зависимости от сезона [100]. Регулярный характер осадков подчёркивает практическую значимость методов восстановления от зашумлений. В настоящей работе внимание сосредоточено на дождевых шумах, устранение которых направлено на повышение устойчивости методов компьютерного зрения в транспортной среде.

Проведенный в 1 главе анализ показал, что атмосферные воздействия приводят к деградации визуальной информации, выражающейся в снижении контрастности, частичной утрате границ объектов, появлении структурных искажений изображения, следствием чего является снижение устойчивости алгоритмов компьютерного зрения. Исследование методов восстановления изображения требует соответствующих данных, однако получение реальных парных наборов данных, содержащих идентичные сцены в зашумлённом и эталонном состоянии, практически невозможно, что ограничивает применение методов восстановления изображения, требующих строго согласованных обучающих выборок. В этих условиях рассматриваются методы, допускающие обучение на синтетических либо несогласованных данных и обеспечивающие преобразование изображений между доменами наблюдения. Особый интерес представляют генеративные методы, реализующие подобное преобразование без необходимости формирования парных обучающих примеров. В то же время влияние уровня восстановления изображения на стабильность последующего анализа дорожной сцены и на потребность в вычислительных ресурсах при обработке видеопотока требует экспериментального исследования. В данной работе рассматриваются представители двух различных подходов к генеративному восстановлению: генеративно-сопоставительные методы (Deraining CycleGAN, SensingGAN) и метод итеративной вероятностной реконструкции (Weather Diffusion).

В качестве первого шага исследования проведено экспериментальное сравнение современных методов восстановления изображений для выбора

базового метода, пригодной для использования в системах потоковой обработки видеоданных в условиях реального времени. Для проведения оценки базовых методов был выбран синтетический набор данных Rain800 [62], содержащий 800 пар изображений с добавленными дождевыми полосами и соответствующие им эталонные изображения без дождя.

Оценка проводилась с применением метрик, отражающих степень сохранения структурных характеристик сцены и уровень искажений восстановленного изображения. Для количественной оценки качества восстановления изображения используются метрики структурного сходства (Structural Similarity Index Measure, SSIM) и отношения максимального значения сигнала к среднеквадратической ошибке (Peak Signal-to-Noise Ratio, PSNR).

Метрика SSIM определяется выражением

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (3.1)$$

где

- x – эталонное изображение;
- y – восстановленное изображение;
- μ_x, μ_y – средние значения яркости изображений;
- σ_x^2, σ_y^2 – дисперсии яркости;
- σ_{xy} – ковариация изображений;
- C_1, C_2 – стабилизирующие константы на базе динамического диапазона пикселей.

Метрика SSIM принимает значения в диапазоне от 0 до 1, при этом большие значения соответствуют более высокому структурному сходству изображений.

Метрика PSNR вычисляется по формуле

$$PSNR = 10 \log_{10} \frac{(L^2)}{MSE}, \quad (3.2)$$

где

- L – максимальное значение, принимаемое пикселем изображения;
- MSE – среднеквадратическая ошибка между эталонным и восстановленным изображениями.

Среднеквадратическая ошибка определяется выражением

$$MSE = \frac{1}{nm} \sum_{i=1}^n \sum_{j=1}^m (x_{ij} - y_{ij})^2, \quad (3.3)$$

где

- n, m – параметры изображения;
- x_{ij} – значения пикселя эталонного изображения;
- y_{ij} – значения соответствующего пикселя восстановленного изображения.

Эксперимент проводился с применением аппаратного графического процессора NVIDIA Tesla A100 с объемом видеопамяти 40 ГБ. Результаты экспериментов, приведенные в таблице 3.1.1, показали, что с практической точки зрения наиболее сбалансированные показатели по качеству и скорости продемонстрировал метод Sensing GAN, вследствие чего он будет рассматриваться как базовый метод. Высокие показатели значений метрик качества восстановления изображения диффузионного метода тем не менее не делают его применимым в задаче обработки видеопотоков за счет низкой скорости обработки изображения.

Таблица 3.1.1 – сравнительная оценка качества восстановления и вычислительной сложности рассматриваемых методов

Метод	PSNR	SSIM	Время обработки изображения, мс
Weather Diffusion	33.1	0.92	12140
Deraining CycleGAN	32.4	0.89	94

Sensing GAN	30.8	0.87	63
-------------	------	------	-----------

Структура генеративных методов допускает снижение емкости нейросетей при сохранении основных принципов реконструкции, позволяя уменьшить вычислительную сложность за счет применения метода дистилляции знаний. Далее проведены исследования возможности снижения емкости рассматриваемых методов и их влияние на качество реконструкции и сохранение объектов дорожной сцены.

3.2. Исследование базовых методов

Исследование влияния дистилляции знаний на качество работы базовых методов. На втором шаге исследования проведена экспериментальная оценка влияния уменьшения емкости генеративных методов на характеристики восстановленного изображения. Снижение сложности реализуется посредством дистилляции знаний, предусматривающей перенос представления учителя в более компактного ученика [101, 102]. Применительно к задачам генеративного восстановления изображений подобное уменьшение параметров может приводить к изменению структуры реконструируемых деталей сцены, что может влиять на корректность последующего анализа динамики объектов, используемых для обнаружения дорожных происшествий.

Основной задачей оценки является определение допустимого уровня уменьшения емкости без существенного ухудшения характеристик восстановленного изображения, влияющих на последующий анализ дорожной сцены. Особое внимание уделяется сравнению поведения различных методов генеративного восстановления при переносе знаний, поскольку различия принципов выполнения вывода нейросети могут приводить к неодинаковой чувствительности к сокращению параметров.

В задаче дистилляции знаний могут применяться различные функции потерь. Однако в задачах генерации изображений для удаления шумов использование только кросс-энтропии недостаточно, так как она не учитывает сохранение высокоуровневых структур и перцептивных характеристик изображения. В таких случаях целесообразно применять перцептуальные функции потерь (Perceptual Loss), основанные на промежуточных признаках предобученных свёрточных сетей, что позволяет нейронной сети сохранять больше информации о содержании изображения и повышает визуальное качество восстановленных данных.

Для проведения дистилляции метода Weather Diffusion в качестве ученика использовался U-Net с уменьшенным вдвое числом каналов в свёрточных слоях относительно родителя: вместо 128 входных каналов в исходной реализации использовалось 64. В качестве функции потерь применялась Perceptual Loss, основанная на признаках, извлекаемых предобученной VGG-19. В качестве нейросети-учителя использовалась предобученная Weather Diffusion, представленная авторами оригинальной работы. Обучение дистиллированной нейронной сети проводилось в течение 100 эпох.

Дистилляция Deraining CycleGAN проводилась по аналогичному принципу. В процессе дистилляции модифицировался генератор, в то время как сегментация капель выполнялась U-ARSE. В роли ученика использовался облегченный генератор с уменьшенным количеством слоев свертки с 18 до 12. В качестве учителя использовалась оригинальная предобученная нейронная сеть.

Для проведения дистилляции метода SensingGAN в качестве ученика использовался облегченный блок удаления дождя на основе U-Net с сокращением количества свёрточных слоёв с 10 до 6. Блок для выделения областей изображения с дождевыми полосами на основе U-Net остался без изменений. В качестве учителя также использовался оригинальный предобученный генератор.

Таблица 3.2.1 – сравнение базовых и дистиллированных методов восстановления

Метод	PSNR	SSIM	Время обработки изображения, мс
Weather Diffusion	33.1	0.92	12140
Weather Diffusion после дистилляции	26.3	0.81	8250
Deraining CycleGAN	32.4	0.89	84
Deraining CycleGAN после дистилляции	31.1	0.83	51
SensingGAN	30.8	0.87	63
SensingGAN после дистилляции	28.4	0.84	41

Полученные в таблице 3.2.1 результаты показывают, что уменьшение емкости нейросетей генеративных методов оказывает различное влияние на характеристики восстановленного изображения и вычислительную сложность выполнения вывода нейросети. Для диффузионного метода сокращение вычислительных затрат приводит к более выраженной потере мелких структурных элементов изображения. Состязательные модели при уменьшении времени обработки кадра лучше сохраняют структуру дорожной сцены, что делает их более подходящей основой для дальнейшего упрощения архитектуры.

Сокращение числа параметров нейронной сети и связанное с этим уменьшение времени выполнения вывода рассматриваются как способ снижения вычислительных затрат на нейросетевой вывод при сохранении характеристик дорожной сцены, используемых на последующих этапах видеоанализа. Следовательно, любое дальнейшее упрощение вычислений должно оцениваться как с точки зрения качества восстановления изображения, так и с учетом степени сохранения структурной целостности. С учетом этого, следующий этап исследования предполагает оценку влияния квантования параметров нейронной

сети на качество реконструкции и сохранение информативных признаков изображения.

Исследование влияния квантизации на качество работы базовых методов. Проведённое выше исследование показало, что снижение емкости генеративных методов посредством дистилляции знаний сопровождается различной степенью деградации качества восстановления изображения для рассматриваемых методов. Вместе с тем вычислительная сложность выполнения вывода нейросети определяется не только количеством параметров нейронной сети, но и точностью их представления. В условиях обработки видеопотоков городской системы наблюдения снижение разрядности параметров может рассматриваться как альтернативный способ уменьшения вычислительной нагрузки, требующий оценки влияния на сохранение информативных признаков дорожной сцены.

Изменение точности представления параметров может приводить к изменению характера восстановленных текстур изображения и границ объектов, что потенциально влияет на устойчивость последующего анализа взаимного расположения транспортных средств и их траекторий. В связи с этим на данном шаге исследования проводится экспериментальная оценка устойчивости генеративных методов восстановления изображения к квантизации параметров.

Снижение точности параметров нейронных сетей выполнялось посредством перехода к представлению весов и активаций с уменьшенной разрядностью. Целью эксперимента являлось определение допустимого уровня квантизации, при котором достигается уменьшение вычислительной сложности обработки кадра без существенного ухудшения характеристик восстановленного изображения.

Один из базовых подходов – пост тренировочная квантизация (PTQ). В этом случае заранее обученная нейронная сеть преобразуется в низкобитное представление без дополнительного обучения. Для уменьшения потерь качества применяется калибровка на небольшом подмножестве данных [103, 104].

Данный метод отличается простотой реализации, однако при сильном снижении разрядности способен существенно ухудшить результаты генерации.

Отдельное направление связано с адаптацией квантизации для диффузионных методов [105, 106]. Поскольку они оперируют в пространстве шума и выполняют последовательные шаги его удаления, предлагается дифференцированный подход: квантизация ранних шагов генерации может быть более грубой, в то время как поздние шаги требуют повышенной точности. Результаты экспериментальной проверки влияния PTQ квантизации на качество работы исследуемых методов представлены в таблице 3.2.2.

Таблица 3.2.2 – результаты сравнительной оценки исходных и квантизированных методов

Метод	PSNR	SSIM	Время обработки изображения, мс
Deraining CycleGAN, FP32	32.4	0.89	84
Deraining CycleGAN, INT8	28.1	0.84	73
Deraining CycleGAN, INT4	23.5	0.73	59
Sensing GAN, FP32	28.4	0.84	64
Sensing GAN, INT8	25.3	0.73	53
Sensing GAN, INT4	21.5	0.65	31
Weather Diffusion, FP32	33.1	0.92	12140
Weather Diffusion, INT8	26.6	0.83	10400
Weather Diffusion, INT4	23.2	0.74	9610

Полученные результаты отражают, что наиболее чувствительным к квантованию оказался диффузионный метод, чье уменьшение точности параметров сопровождается потерей мелких деталей и изменением текстуры изображения. Методы на основе GAN сохраняют структуру сцены устойчивее, при этом сокращение времени нейросетевого вывода остается сопоставимым.

Следует отметить, что уменьшение точности параметров приводит не только к изменению визуальных характеристик восстановленного изображения, но и может оказывать влияние на различимость границ транспортных средств и элементов дорожной инфраструктуры. Подобные изменения потенциально способны снижать устойчивость последующего анализа дорожной сцены и определения происшествий.

Полученные результаты выявили дополнительное ограничение, связанное с устойчивостью генеративных методов к изменению точности параметров, что требует дальнейшей оценки влияния восстановления изображения на задачи анализа дорожной сцены. В связи с этим на следующем этапе исследования проводится экспериментальная оценка влияния восстановления изображения на качество обнаружения объектов дорожной сцены.

Исследование влияния методов удаления шумов с изображений на качество работы алгоритмов детекции. Проведённые ранее исследования позволили определить особенности поведения генеративных методов восстановления изображения при снижении вычислительной сложности посредством уменьшения емкости нейронных сетей и точности представления параметров. Вместе с тем оценка выполнялась преимущественно по метрикам качества реконструкции изображения и времени обработки одного кадра, которые не позволяют непосредственно определить влияние восстановления изображения на устойчивость обнаружения ДТП.

В связи с этим на данном шаге исследования проводится оценка влияния атмосферных деградаций и их последующего устранения на качество обнаружения объектов дорожной сцены.

В качестве инструмента детекции используется алгоритм YOLOv8, включённый в состав разработанного во второй главе метода обнаружения ДТП и обеспечивающая локализацию объектов транспортной сцены на этапе формирования признаков.

Оценка выполняется по метрике средней точности обнаружения (mAP), позволяющей количественно оценить изменение наблюдаемости объектов

дорожной сцены при различных условиях визуальной деградации изображения. Преимущество данной метрики заключается в том, что она отражает способность метода корректно определять местоположение объектов разных категорий и минимизировать количество ложных срабатываний, что особенно важно в задаче анализа дорожной сцены, где ошибки распознавания отдельных классов могут в значительной степени повлиять на итоговую оценку ситуации. Формула для нахождения mAP:

$$mAP = \frac{1}{n_{classes}} \times \sum_{c \in classes} \frac{TP_c}{TP_c + FP_c}, \quad (3.4)$$

где $n_classes$ – количество классов, $classes$ – множество всех классов, TP_c – количество истинно положительных прогнозов для класса c , FP_c – количество ложно положительных прогнозов для класса c .

Перед оценкой эффективности методов восстановления была проведена сравнительная оценка качества обнаружения объектов дорожной сцены в условиях ясной погоды и при наличии дождевых артефактов. В качестве данных для эксперимента была выбрана валидационная выборка набора данных BDD100k [107], которая содержит 10000 изображений размеченных объектов 12 различных классов, включая автомобили, дорожный знак, светофор, человек.

Цель данного эксперимента заключается в количественном подтверждении влияния атмосферных искажений на наблюдаемость объектов дорожной сцены в рамках используемого метода обнаружения дорожно-транспортных происшествий. Результаты эксперимента представлены в таблице 3.2.3.

Таблица 3.2.3 – качество обнаружения объектов дорожной сцены при различных погодных условиях (mAP)

	Car	Person	Traffic light	Traffic Sign	All
Ясная погода	0.54	0.36	0.39	0.3	0.4
Дождь	0.33	0.13	0.19	0.15	0.2

Полученные результаты показывают, что наличие дождевых артефактов приводит к снижению качества обнаружения объектов по всем рассматриваемым классам. Наиболее выраженное снижение наблюдается для объектов малых размеров, включая пешеходов и дорожные знаки, для которых ухудшение различимости границ оказывает наибольшее влияние.

Проведена оценка влияния генеративного восстановления изображений на качество детекции. Для этого выполнялась обработка изображений с дождевыми искажениями всеми исследуемыми методами с последующей оценкой качества обнаружения объектов посредством применения алгоритма YOLO.

Таблица 3.2.4 – качество обнаружения объектов дорожной сцены после восстановления рассматриваемыми методами (mAP)

	Car	Person	Traffic light	Traffic Sign	All
Weather Diffusion	0.52	0.33	0.36	0.28	0.36
Derain CycleGAN	0.48	0.25	0.36	0.27	0.36
SensingGAN	0.51	0.31	0.35	0.25	0.36

Данные, представленные в таблице 3.2.4, свидетельствуют о том, что восстановление изображений с использованием генеративных методов повышает показатель mAP по сравнению с зашумленными данными. Сравнение с результатами, приведенными в таблице 3.2.3, показывает, что сам факт устранения артефактов, вызванных дождем, оказывает более существенное влияние на качество обнаружения, чем различия между рассматриваемыми методами восстановления.

3.3. Разработка метода отчистки изображения от погодных шумов

Особенности реализации метода обработки входных данных предполагают одновременную работу с несколькими источниками входных данных. Поскольку

восстановление изображения в условиях, вызванных атмосферными осадками, осуществляется на протяжении всего периода затрудненной видимости, вычислительная сложность этого этапа возрастает пропорционально количеству видеопотоков и частоте кадров. Дополнительная сложность задачи заключается в том, что качество обнаружения ДТП зависит от точности восстановления геометрических и структурных характеристик дорожной сцены. Глобальные преобразования изображения могут исказить неизменные участки сцены, содержащие элементы дорожной инфраструктуры, тем самым снижая надежность последующего анализа взаимного расположения транспортных средств.

Разработка метода восстановления. В отличие от рассмотренных методов, выполняющих глобальное преобразование изображения, в SensingGAN использование карты дождевых признаков позволяет ограничить область изменения изображения деградированными участками сцены и снизить риск искажения неизменённых областей. При обработке множественных видеопотоков одновременно существенное влияние на производительность оказывает параметрическая емкость генератора и объем выполняемых вычислений. В связи с этим в рамках диссертационной работы разработан компактный генератор, сохраняющий принцип обработки изображения базового метода, но предназначенный для последующего обучения методом дистилляции знаний и адаптированный для потоковой обработки.

Входной модуль выполняет преобразование четырехканального входного тензора в пространство признаков размерностью 64 канала посредством свертки с ядром 5×5 , позволяющей извлекать локальные структурные особенности изображения и карты дождевых признаков при сохранении пространственного разрешения.

Кодировщик состоит из двух последовательных уровней уменьшения пространственного разрешения. Каждый уровень включает сверточный слой, обеспечивающий снижение размеров карты признаков и увеличением числа

каналов с 64 до 256, формируя компактное описание изображения с локальными признаками текстур, границ объектов и областей с дождевыми искажениями.

Блок самовнимания предназначен для моделирования зависимостей между удаленными в пространстве областями изображения и формирования глобальных признаков.

Модуль обработки признаков разного масштаба включает две последовательно выполняемые дилатированные свертки с коэффициентами расширения 2 и 4, обеспечивающие последовательное увеличение эффективного поля восприятия без уменьшения пространственного разрешения карт признаков и учитывающие разномасштабные структуры.

Декодировщик построен симметрично кодировщику и последовательно восстанавливает пространственное разрешение на двух уровнях. На каждом из них производится увеличение карт признаков выполняется транспонированной сверткой, после чего полученные признаки объединяются с данными соответствующего уровня кодировщика через пропускающее соединение, сохраняя детали, теряемые при уменьшении разрешения.

После объединения признаков на каждом уровне выполняется дополнительная сверточная обработка, уточняющая восстановленное представление. На завершающем этапе используется выходной сверточный слой, преобразующий внутреннее представление в трехканальное восстановленное изображение.

По сравнению с исходной архитектурой число обучаемых параметров генератора уменьшено с 6.1 до 3.8 миллионов параметров, что соответствует сокращению на 38 %. Достигнутое уменьшение вычислительной сложности обеспечивает возможность последующего применения процедур дистилляции знаний и целочисленной квантизации без изменения структуры сети.

Шаг 1. Предварительная обработка входного изображения. Входное изображение $I \in \mathbb{R}^{H \times W \times 3}$, где H, W - высота и ширина изображения, а три канала соответствуют цветовой модели RGB, приводится к формату, используемому генератором. Выполняется изменение пространственного разрешения до

фиксированного размера, преобразование в тензор и нормализация значений пикселей. В результате формируется нормализованное изображение:

$$I_p = P(I), \quad (3.5)$$

где $P(\cdot)$ – оператор предварительной обработки, включающий масштабирование изображения, преобразование в тензор и нормализацию.

Шаг 2. Формирование карты дождевых признаков. По предварительно обработанному изображению вычисляется карта признаков дождевых артефактов:

$$R = F(I_p), \quad (3.6)$$

где $F(\cdot)$ – оператор построения карты дождевых признаков.

Полученная карта выделяет области изображения, содержащие характерные пространственные структуры атмосферных осадков.

Шаг 3. Формирование входного представления генератора. Нормализованное изображение и карта дождевых признаков объединяются по каналному измерению:

$$X = \text{Concat}(I_p, R), \quad (3.7)$$

где $X \in \mathbb{R}^{H \times W \times 4}$ – выходной четырехканальный тензор.

Шаг 4. Извлечение пространственных признаков. Из входного тензора X происходит извлечение признаков посредством сверточных операций с одновременным уменьшением пространственного разрешения.

$$F_1 = \sigma \left(C_{5 \times 5, S=1}^{64}(X) \right), \quad (3.8)$$

$$F_2 = \sigma \left(C_{3 \times 3, S=2}^{128}(F_1) \right), \quad (3.9)$$

$$F_3 = \sigma \left(C_{3 \times 3, S=1}^{128}(F_2) \right), \quad (3.10)$$

$$F_4 = \sigma \left(C_{3 \times 3, S=2}^{256}(F_3) \right), \quad (3.11)$$

$$F_5 = \sigma \left(C_{3 \times 3, S=1}^{256}(F_4) \right), \quad (3.12)$$

$$F_6 = \sigma \left(C_{3 \times 3}^{256}(F_5) \right), \quad (3.13)$$

где

- F_c – последовательные карты признаков;
- $C_{x \times x, s=m}^n$ – операция двумерной свёртки с размером ядра $x \times x$, количеством выходных каналов n , $s = m$ – шагом смещения ядра свертки;
- $\sigma(\cdot)$ – функция активации ReLU.

В результате выполнения этапа формируется компактное признаковое представление входного изображения.

Шаг 5. Обработка признакового пространства. Выполняется обработка сформированного признакового представления с использованием механизма самовнимания и дилатированных свёрток:

$$F_7 = SA(F_6), \quad (3.14)$$

$$F_8 = \sigma \left(C_{3 \times 3, d=2}^{256}(F_7) \right), \quad (3.15)$$

$$F_9 = \sigma \left(C_{3 \times 3, d=4}^{256}(F_8) \right), \quad (3.16)$$

где

- $SA(\cdot)$ – модуль самовнимания, обеспечивающий учет глобальных пространственных зависимостей между признаками;
- $C_{x \times x, d=m}^n$ – операция двумерной дилатированной свёртки с размером ядра $x \times x$, количеством выходных каналов n , $s = m$ – коэффициентом дилатации.

В результате выполнения этапа формируется признаковое представление, содержащее как локальные, так и глобальные характеристики изображения, необходимые для последующего восстановления.

Шаг 6. Восстановление изображения. Преобразованные признаки поступают в декодирующую часть генератора, восстанавливающую пространственное разрешение изображения:

$$F_{10} = U_{256 \rightarrow 128}(F_9) + F_3, \quad (3.17)$$

$$F_{11} = \sigma \left(C_{3 \times 3, s=1}^{128}(F_{10}) \right), \quad (3.18)$$

$$F_{12} = U_{128 \rightarrow 64}(F_{11}) + F_1, \quad (3.19)$$

$$F_{13} = \sigma \left(C_{3 \times 3, s=1}^{32}(F_{12}) \right), \quad (3.20)$$

где $U_{k \rightarrow j}(\cdot)$ – блок увеличения пространственного разрешения в два раза с уменьшением числа каналов с k до j .

В результате выполнения этапа формируется полноразмерное признаковое представление, содержащее восстановленные пространственные детали и используемое для формирования трёхканального выходного изображения.

Шаг 7. Формирование выходного изображения. Выполняется восстановление цветового изображения на основе сформированного признакового представления:

$$I = C_{3 \times 3}^3(F_{13}), \quad (3.21)$$

где $\hat{I}$ – восстановленное изображение.

После формирования выходного изображения выполняется удаление дополненных при предварительной обработке областей, что обеспечивает соответствие результата исходному размеру изображения.

Обучение разработанной нейросети. В данной работе применялся один из подходов к обучению в задачах сжатия нейронных сетей, при котором нейросеть-ученик аппроксимирует отображение, формируемое предварительно обученной нейросетью-учителем. При этом учитель и ученик могут быть построены на одной архитектурной основе, отличаясь глубиной сети, количеством каналов или числом параметров. В качестве модели-учителя используется предобученный генератор базового метода. После загрузки его параметры фиксируются и в процессе дистилляции не изменяются. В качестве модели-ученика используется разработанный облегчённый генератор, параметры которого подлежат оптимизации. На каждой итерации обе модели получают одинаковые входные изображения, после чего формируются восстановленные изображения учителя и ученика. На основе полученных результатов вычисляется совокупная функция потерь, включающая сумму Perceptual Loss, GANLoss и SSIMLoss.

Для обучения было использовано 5000 изображений из набора данных BDD100k с наложенными на них характерными синтетическими шумами на протяжении 250 эпох. Оптимизация параметров модели-ученика выполняется с использованием алгоритма AdamW, размер батча 8.

Полученные максимальные валидационные значения $PSNR = 24,7$ и $SSIM = 0,9$ свидетельствуют о высокой степени сохранения структурного сходства восстановленного изображения с эталонным.

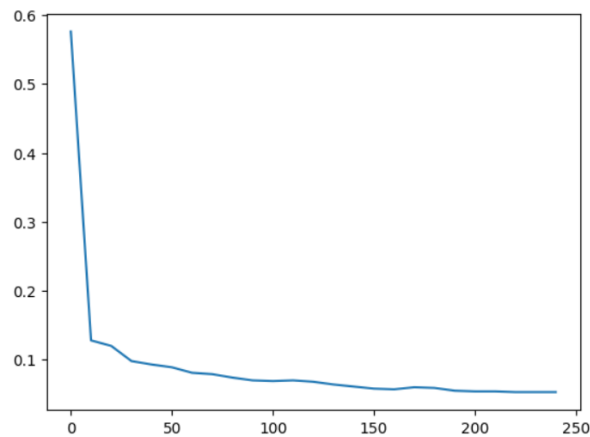


Рисунок 3.1 – График Loss при обучении разработанного генератора

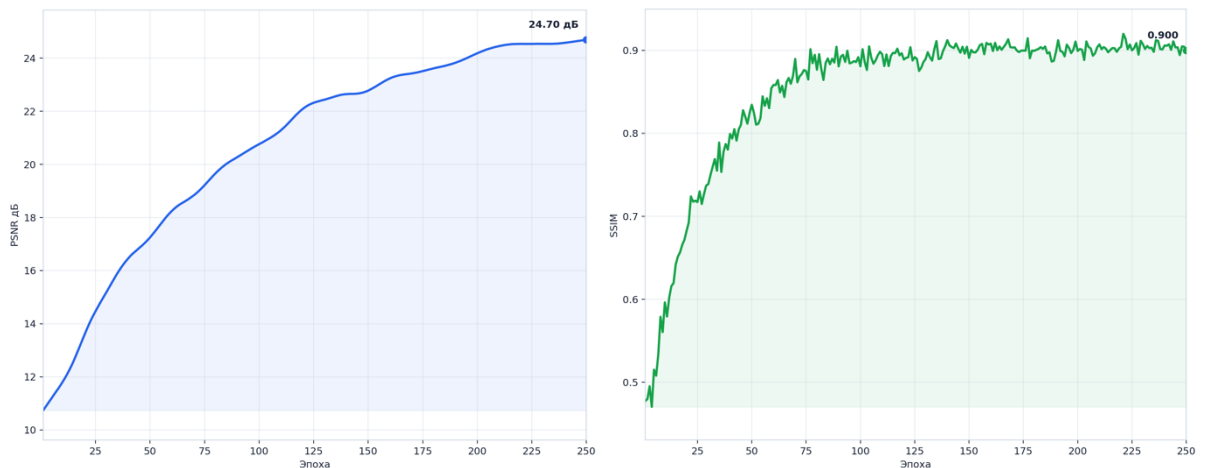


Рисунок 3.2 – Валидационные графики PSNR (слева) и SSIM (справа) при обучении разработанного генератора

После завершения обучения разработанного генератора выполняется его посттренировочная статическая INT8-квантизация средствами TensorRT. Для определения параметров квантования используется калибровочная выборка

объёмом 300 изображений из тестовой выборки набора данных BDD100k. По результатам квантизации значения метрик достигли $PSNR = 27.4$, $SSIM = 0.78$, а скорость обработки кадра составила 29 мс и превысила скорость базового метода в два раза.

Для оценки применимости разработанного метода проведен анализ влияния восстановления изображения на качество работы алгоритма YOLOv8m. Результаты представлены в таблице 3.3.1.

Таблица 3.3.1 – влияние удаления дождя на качество работы YOLO по метрике mAP

	Car	Person	Traffic light	Traffic sign	All
Ясная погода	0.54	0.36	0.39	0.3	0.4
Дождь	0.33	0.13	0.19	0.15	0.2
Удаление дождя базовым Sensing GAN	0.51	0.31	0.35	0.25	0.36
Удаление дождя разработанным методом	0.53	0.3	0.34	0.24	0.35

Анализ влияния восстановления изображения на работу алгоритма детекции показал, что при наличии дождевых деградаций среднее значение mAP по всем классам снижается до 0.2. Использование метода с модифицированной конфигурацией генератора обеспечивает значение mAP 0,35, что сопоставимо с результатом базового метода в 0.36 при существенном снижении вычислительной сложности. Наблюдаемое отличие по отдельным классам не приводит к ухудшению общей эффективности обнаружения по сравнению с базовым методом.

Сопоставление значений SSIM и mAP показывает, что уменьшение параметрической емкости генератора не привело к существенному снижению качества последующего обнаружения объектов. Это подтверждает, что сохраняются механизмы передачи пространственной информации, определяющие возможность последующего анализа изображения.

На рисунке 3.2, взятом из тестовой выборки набора данных BDD100k, наглядно иллюстрировано влияние разработанного метода восстановления изображения на последующее обнаружение объектов алгоритмом YOLOv8m.



Рисунок 3.2 – Влияние восстановления изображения от дождевых деградаций на результат обнаружения объектов дорожной сцены: слева – исходное изображение с дождевыми артефактами и результат детекции; справа – изображение после восстановления разработанным методом и результат детекции.

После восстановления уменьшается визуальная плотность дождевых полос, растет общий порог уверенности прогнозов классов обнаруженных объектов, а также наблюдается главное следствие восстановления изображения – улучшается детекция автомобилей, ранее не обнаруженных на зашумленных данных из-за искажения визуальных признаков.

Полученные результаты демонстрируют, что снижение сложности генератора при сохранении механизмов локализации дождевых деградаций обеспечивает восстановление структурных признаков дорожной сцены на уровне, достаточном для поддержания эффективности задачи анализа объектов дорожной сцены, и позволяет интегрировать метод восстановления изображения в систему обнаружения дорожно-транспортных происшествий.

3.4. Выводы по главе

В третьей главе выполнено исследование современных методов восстановления изображений, искажённых атмосферными осадками. Установлено, что существующие методы характеризуются различным соотношением качества восстановления и вычислительной сложности.

Показано, что существующие методы обладают ограниченной способностью локализовать области атмосферных искажений, что приводит к снижению качества восстановления при сохранении вычислительной эффективности.

Разработан метод восстановления изображения от погодных шумов, основанный на совместном использовании карты дождевых признаков, облегчённого генератора. Число обучаемых параметров генератора уменьшено с 6,1 до 3,8 млн, что соответствует сокращению на 38 % по сравнению с исходной архитектурой.

После обучения генератора значение структурного сходства восстановленного изображения сохраняется в значении не менее 0,78, а время обработки одного кадра – не более 29 мс, что соответствует двукратному сокращению времени обработки по сравнению с базовым методом.

Экспериментальная оценка показала, что после восстановления изображений с дождевыми искажениями с применением разработанного метода среднее значение mAP при обнаружении объектов увеличивается с 0,20 до 0,35.

Разработанный метод восстановления изображения обеспечивает сочетание структурного сходства восстановленного изображения, вычислительной эффективности и сохранения информативности визуальных признаков, необходимые для применения в составе автоматизированной информационной системы обнаружения дорожно-транспортных происшествий в условиях затруднённой видимости.

Глава 4. Разработка АИС для обнаружения ДТП в затрудненных погодных условиях

В четвертой главе решается задача разработки архитектуры автоматизированной информационной системы обнаружения дорожно-транспортных происшествий в условиях затрудненной видимости. Разрабатываемая АИС должна обеспечивать полный цикл обработки информации, включая получение видеопотоков от камер дорожного наблюдения, восстановление изображений, обнаружение и анализ объектов дорожной сцены, формирование событий ДТП, хранение результатов и отправку во внешние информационные системы.

При проектировании АИС учитываются различия в вычислительных затратах, характерные для разных этапов нейросетевой обработки. Архитектура системы отличается модульностью, масштабируемостью и возможностью независимого обновления компонентов, а также предусматривает разделение оперативных данных и информации, предназначенной для длительного хранения. Описаны составные компоненты АИС и взаимосвязи между ними, процессы обработки видеопотока и передачи данных между компонентами, а также процедуры формирования и хранения информации о ДТП. В завершение представлена информационная модель системы и проведена экспериментальная оценка ее эффективности.

Разработанные в главах 2 и 3 нейросетевые методы обнаружения ДТП и восстановления изображений интегрированы в АИС. Их совместная работа предполагает выполнение вычислений в определенной последовательности, адаптивный выбор режимов обработки в зависимости от состояния входного видеопотока и автоматизированное распределение вычислительных ресурсов между решаемыми задачами.

4.1. Архитектура АИС

Выбор архитектурного подхода к построению АИС определяется совокупностью требований, а также результатами исследований методов восстановления изображений и детекции ДТП, представленных в главах 2 и 3.

Функциональные требования к АИС включают прием и декодирование видеопотоков, оценку погодных условий, восстановление изображений, анализ движения транспортных средств, а также обнаружение ДТП с сохранением соответствующих данных. Архитектура системы также должна обеспечивать масштабируемость, независимую работу компонентов и контролируемое использование вычислительных ресурсов.

С учетом этих требований функции АИС распределены между автономными компонентами, имеющими четко определенные интерфейсы взаимодействия. Поскольку монолитная реализация препятствует независимому масштабированию и обновлению наиболее ресурсоемких этапов, в основе АИС лежит микросервисная архитектура, представленная на рисунке 4.1.

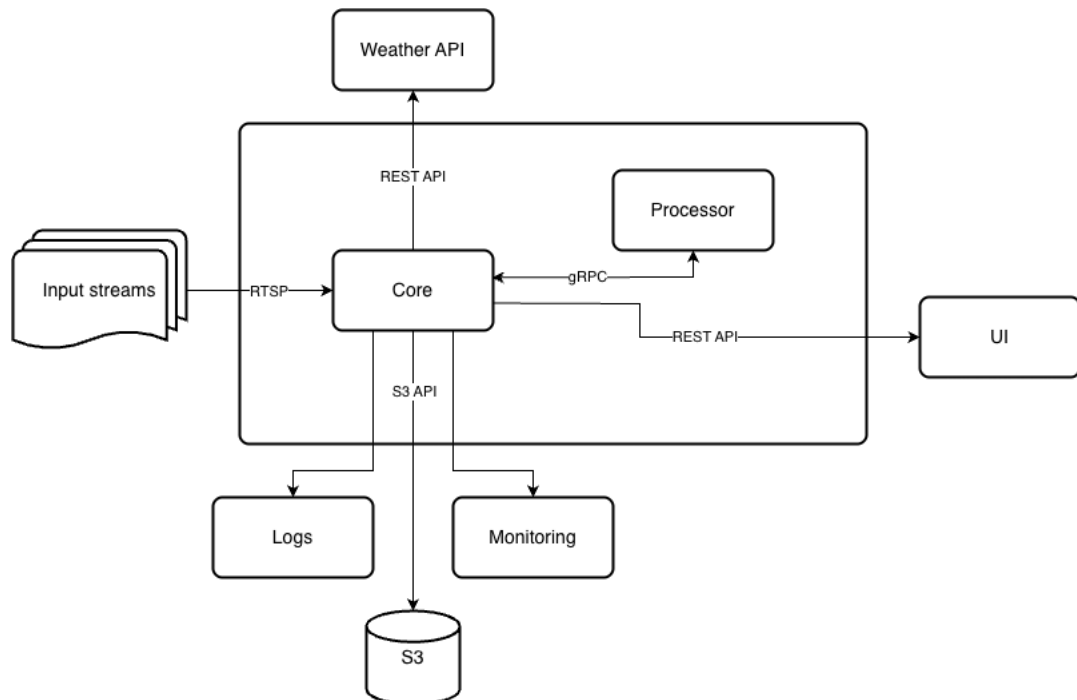


Рисунок 4.1 – Архитектура разрабатываемой системы

Архитектура разработанной АИС включает центральный управляющий компонент Core, вычислительный компонент Processor, компоненты журналирования Logs и мониторинга Monitoring, объектное хранилище S3, а также внешние источники и потребители данных. Центральным элементом архитектуры является компонент Core, координирующий взаимодействие остальных компонентов системы, тогда как выполнение вычислительно сложных операций вынесено в компонент Processor, что позволяет независимо масштабировать вычислительные ресурсы. Для долгосрочного хранения сформированных видеоклипов обнаруженных происшествий добавлено объектное хранилище S3. Компоненты Logs и Monitoring обеспечивают эксплуатационное сопровождение системы. Взаимодействие с внешней средой осуществляется посредством входных видеопотоков (Input Streams), сервиса погодных данных (Weather API) и интерфейса передачи результатов (Output). Подробное описание назначения, принципов функционирования и взаимодействия каждого компонента приведено в последующих подразделах главы.

Разрабатываемая АИС рассчитана на серверное развертывание с возможностью горизонтального масштабирования. Обработка видеоданных выполняется на одном или нескольких серверах, что обеспечивает централизованное управление вычислительными ресурсами.

В процессе функционирования автоматизированной информационной системы используются информационные объекты, переходящие между архитектурными слоями обработки видеопотока и определяющие структуру их взаимодействия. К таким объектам относятся конфигурационные данные источников видеонаблюдения, кадровые данные видеопотока, результаты нейросетевой обработки изображений, а также событийные данные, формируемые при обнаружении ДТП. Особенности их представления, хранения и передачи между процессами рассматриваются в последующих разделах главы.

4.2. Функциональная структура АИС

Функциональная структура АИС ориентирована на обработку непрерывных видеопотоков, поступающих с множественных камер. Структура обработки организована в виде набора логических слоёв, каждый из которых решает строго определённую задачу и взаимодействует с другими слоями через интерфейсы. Слоистая структура позволяет изолировать вычислительно сложные этапы обработки данных и обеспечить независимую масштабируемость компонентов. Схема типовой части функциональной структуры и взаимодействия ее компонентов представлена на рисунке 4.2.

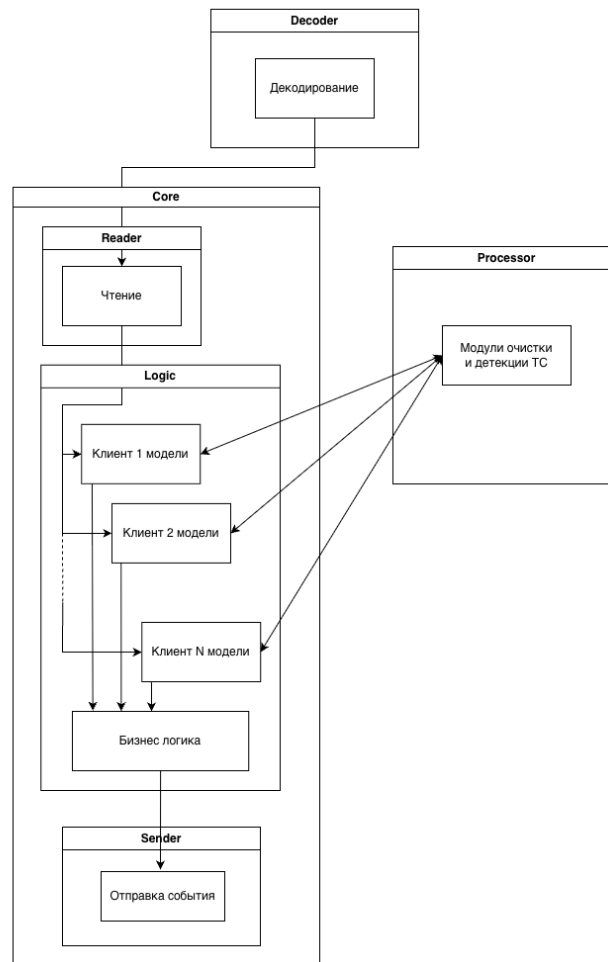


Рисунок 4.2 – Функциональная схема взаимодействия Core

Первый этап обработки видеоданных в АИС реализован в отдельном слое Decoder. Он реализует захват и декодирование видеопотоков, поступающих от

камер дорожного видеонаблюдения с использованием аппаратного декодировщика. Decoder выполняет преобразование входных RTSP-потоков в последовательность декодированных кадров и размещает их в области общей памяти (shared memory). Использование общей памяти в качестве промежуточного хранилища данных позволяет отделить процесс декодирования видеопотока от последующих этапов анализа изображений и исключить необходимость передачи больших объёмов данных через межпроцессные каналы общего назначения. Такое преобразование обеспечивает устранение избыточного копирования кадров между процессами, снижая нагрузку на центральный процессор и шину памяти [108], и дает доступ к одному и тому же кадру нескольким потребителям данных без повторного декодирования видеопотока, что особенно важно при параллельной обработке изображений различными подсистемами АИС.

При проектировании слоя захвата и декодирования видеопотока учитывались ограничения, связанные с масштабированием системы по числу источников видеоданных. В условиях обработки множества RTSP-потоков наивная организация видеоввода, предполагающая выделение отдельного процесса декодирования для каждого источника, приводит к избыточному потреблению вычислительных ресурсов, в том числе видеопамати при использовании аппаратного декодирования. В связи с этим Decoder реализован как централизованный компонент, обеспечивающий унифицированный видеоввод и изоляцию процессов декодирования от логики последующей обработки видеоданных.

Следующий слой Reader реализует функцию чтения кадров из общей памяти и их маршрутизации между вычислительными компонентами системы. Его внедрение позволяет разделить операции доступа к видеоданным и непосредственное выполнение методов нейронных сетей. Компонент Reader не выполняет ресурсоемких вычислений, а определяет дальнейший маршрут кадра и направляет его соответствующему компоненту, так как кадры из нескольких видеопотоков могут одновременно находиться на разных стадиях обработки.

Центральным компонентом функциональной структуры АИС является слой Logic, в котором реализуется бизнес-логика анализа дорожной сцены и принятия решений о наличии ДТП. Логический слой принимает кадровые данные видеопотока, а также результаты нейросетевой обработки, агрегирует контекстную информацию и формирует управляющие воздействия для вычислительных компонентов системы. Logic поддерживает скользящее временное окно кадров фиксированной длительности, используемое для реализации метода обнаружения ДТП, описанного в главе 2. Указанное временное окно служит источником данных для формирования и сохранения видеоклипа происшествия при подтверждении факта ДТП.

В слое Logic реализован механизм определения необходимости применения разработанного метода восстановления изображения. Конфигурация каждой камеры содержит сведения о ее местоположении, которые используются для запроса актуальных метеорологических данных у внешних погодных сервисов через программный интерфейс. На основании полученной информации оцениваются условия наблюдения для соответствующего видеопотока и принимается решение о включении нейросетевого восстановления. При наличии погодных факторов, способных ухудшить качество изображения, кадры направляются на обработку методом, разработанным в третьей главе. В остальных случаях этот этап пропускается, и дальнейший анализ выполняется по исходным кадрам. За счет адаптивного включения нейросетевого вывода метода восстановления исключается постоянное выполнение ресурсоемких операций, изображение не проходит нетребуемый этап, понижающий структурную целостность, и вычислительные ресурсы используются только при наличии соответствующих условий наблюдения.

Вычисления с использованием нейронных сетей вынесены на специализированный вычислительный слой Processor, предназначенный для выполнения логики вывода на основе разработанных нейросетевых методов по запросам от слоя бизнес-логики. В свою очередь Processor не содержит бизнес-

логики и выполняет исключительно функции управления экземплярами нейросетевых методов и распределения вычислительной нагрузки.

Для повышения эффективности использования вычислительных ресурсов на слое Processor реализована возможность совместного использования несколькими видеопотоками экземпляров нейронных сетей. В рамках организации этой процедуры при настройке процесса обработки Logic запрашивает нейросеть с определенной конфигурацией. Если экземпляр такой модели уже загружен и позволяет отправлять в одном пакете изображения с камеры, новый поток закрепляется за ним. В случае отсутствия доступного экземпляра процессор создает новый и добавляет его в пул вычислительных ресурсов.

После обнаружения факта ДТП слой Logic формирует событие и передает его слою отправки Sender. На этом этапе подготавливаются сообщения, а информация об инциденте передается во внешние системы посредством программного интерфейса.

Для обеспечения надёжности и управляемости системы в структуре АИС предусмотрено логирование процессов и мониторинг состояния АИС и эксплуатируемых ресурсов, реализованное в слое Monitoring.

АИС рассчитана на одновременную обработку независимых видеопотоков, которые могут различаться частотой кадров, качеством изображения и интенсивностью регистрируемых событий. При синхронном взаимодействии компонентов задержка на одном из этапов способна приостановить последующие операции и привести к накоплению необработанных данных. Для распределения поступающей нагрузки между доступными вычислительными ресурсами и сохранением работоспособности системы при изменении числа одновременно обрабатываемых потоков обмен между компонентами АИС организован асинхронно. Отдельные операции выполняются параллельно, а завершение одного вычислительного процесса не блокирует обработку данных, не связанных с ним.

Архитектурные слои при этом функционируют независимо и обмениваются данными через установленные интерфейсы. Наиболее явно данный принцип проявляется при взаимодействии слоев Logic и Processor. Слой бизнес-логики формирует запрос на нейросетевую обработку кадра и передает его в Processor, продолжая работу с поступающими данными. Processor выполняет нейросетевой вывод по мере освобождения вычислительных ресурсов и возвращает полученный результат в Logic для дальнейшей обработки. Благодаря этому скорость выполнения нейросетевых операций не задает жесткую последовательность работы остальных компонентов системы. В результате латентность обработки определяется наиболее длительной операцией, а не суммой времён выполнения всех этапов анализа. При этом принятие решения о наличии ДТП остается синхронным процессом. Для каждого видеопотока кадры обрабатываются в определенном порядке, зависящим от текущего состояния и ранее полученных результатов выполнения вывода нейросетей.

Состояние каждого видеопотока ведется в Logic независимо от других источников. Для него сохраняются временные окна наблюдений и агрегированные признаки, необходимые для обнаружения ДТП. Разделение состояния потоков и асинхронное выполнение вычислений позволяют локализовать задержки в пределах отдельных операций и не распространять их на обработку остальных камер.

Основная вычислительная нагрузка при работе АИС обусловлена обработкой кадров с помощью нейронных сетей, включающей обнаружение объектов в дорожной сцене и восстановление изображений при наличии атмосферных осадков. Если бы для каждого видеопотока выделялся отдельный экземпляр соответствующего нейросетевого метода, количество загруженных моделей росло бы пропорционально числу подключенных камер. Однако вычислительные мощности одного экземпляра могли бы использоваться неэффективно, поскольку кадры поступают независимо из разных источников, в

то время как нейросетевой механизм вывода (инференса) способен обрабатывать запросы от нескольких видеопотоков одновременно. В связи с этим предложенная архитектура обеспечивает независимость экземпляров нейросетевых методов от конкретных видеопотоков, позволяя динамически распределять их с учетом текущей вычислительной нагрузки.

Для устранения указанного ограничения в разработанной АИС нейросетевые вычисления вынесены в специализированный вычислительный слой Processor, реализующий централизованное управление экземплярами нейросетевых методов и распределение вычислительной нагрузки между ними. Логический слой системы не взаимодействует с конкретными экземплярами нейронных сетей напрямую, а формирует запросы слою Processor на выполнение обработки, в том числе указывая о наличии или отсутствии необходимости восстановления кадра от зашумлений с последующей детекцией транспортных средств. Выполнение нейросетевого вывода происходит асинхронно по отношению к логическому слою, возвращая результаты обработки по завершении без блокировки работы логического слоя.

Для повышения эффективности использования вычислительных ресурсов в слое Processor реализован механизм пула нейронных сетей. Пул представляет собой совокупность активных экземпляров, каждый из которых может обслуживать несколько независимых запросов на выполнение вывода нейросетей. Каждый экземпляр в пуле характеризуется набором параметров, включающих тип, конфигурацию и допустимое число параллельных запросов, определяемое количеством доступных вычислительных слотов, определяющих способность экземпляра обрабатывать несколько кадров параллельно без потери данных.

При необходимости выполнения нейросетевой обработки логический слой формирует запрос в вычислительный слой Processor, указывая требуемый тип нейронной сети и параметры обработки. Processor анализирует текущее состояние пула и выполняет выбор экземпляра в соответствии со следующим алгоритмом:

- производится поиск активного экземпляра с заданными параметрами;
- при наличии экземпляра с доступными вычислительными слотами запрос ассоциируется с данным экземпляром;
- при отсутствии свободных слотов создаётся новый экземпляр, который добавляется в пул и используется для обработки запроса.

Архитектура вычислительного слоя Processor обеспечивает эффективность всей системы за счет повышения пропускной способности системы при обработке множества видеопотоков и является важным элементом масштабируемости АИС.

4.3. Организация взаимодействия процессов в АИС

Функционирование АИС предполагает взаимодействие между несколькими независимыми процессами, реализующими различные архитектурные слои обработки видеопотока. При этом основную нагрузку на подсистему передачи данных формируют видеокадры, представляющие собой высокоразмерные массивы данных. Передача таких данных между процессами в виде сериализованных сообщений может привести к росту латентности обработки, что делает необходимым использование специализированных механизмов межпроцессного взаимодействия. На рисунке 4.3 представлена диаграмма жизненного цикла информационных объектов АИС, показывающая разделение обработки на поток передачи высокоразмерных видеоданных и поток управляющих метаданных.

Обработка видеопотока в АИС распределена между несколькими независимо функционирующими процессами, соответствующими различным архитектурным слоям. Изображения, представляющие собой многомерные массивы значений пикселей, составляют основной объем передаваемых данных. Последовательная сериализация и передача этих кадров совместно с

управляющими сообщениями привели бы к увеличению объема межпроцессного взаимодействия и возникновению дополнительных задержек при обработке. В связи с этим в АИС разделены механизмы обработки видеоданных и механизмы управления сопутствующими метаданными. Жизненный цикл соответствующих информационных объектов показан на рисунке 4.3, где отдельно представлены перемещение кадров между этапами обработки и обмен управляющей информацией между компонентами АИС.

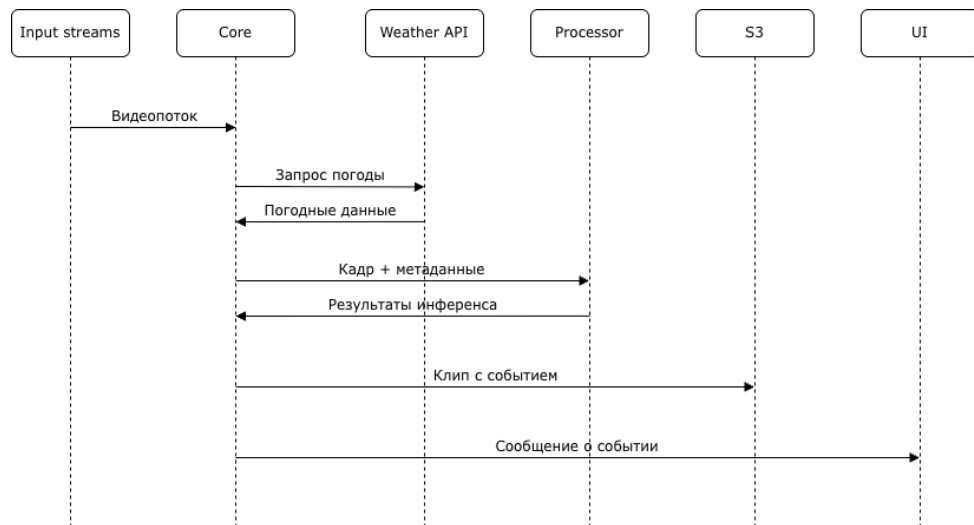


Рисунок 4.3 – UML-диаграмма взаимодействия компонентов АИС при обработке данных

В разработанной АИС обмен данными между слоями организован таким образом, чтобы отделять видеоданные от сопутствующих метаданных. Декодированные кадры не передаются напрямую между процессами, вместо этого они сохраняются в разделяемой памяти, а компонентам системы передается лишь ссылка на соответствующую область памяти. Метаданные, содержащие информацию, необходимую для идентификации кадра и его обработки на последующих этапах, передаются отдельными сообщениями.

Разделяемая память организована в виде набора буферов фиксированного размера, позволяя различным компонентам обращаться к одному и тому же кадру без создания дополнительных копий или передачи всего массива пиксельных данных при каждом межпроцессном взаимодействии. Чтение из буферов не блокирует другие процессы, что дает возможность параллельно

использовать один и тот же кадр на разных этапах обработки. В результате сокращается объем операций копирования видеоданных и устраняется необходимость в последовательной передаче кадров между компонентами AIS.

Между слоями обработки передаются не фактически изображения, а ссылки на соответствующие области общей памяти и сопутствующие метаданные. Непосредственно ссылкой на видеоданные представляется идентификатор области общей памяти и смещение в ней, достаточные для однозначного доступа к соответствующему кадру без передачи самих пиксельных данных.

При взаимодействии между слоями Decoder выполняет захват и декодирование RTSP-видеопотоков и размещает полученные кадры в общей памяти. Reader осуществляет чтение ссылок на размещённые кадры и формирует сообщения с метаданными, которые передаются в Logic. В зависимости от текущего состояния и контекста обработки, последний формирует запросы на выполнения вывода нейросетей получаемых кадров и передаёт их в Processor. Взаимодействие между логическим и вычислительным слоями реализуется в виде асинхронных удалённых вызовов, передающих метаданные запроса и ссылки на видеоданные, при этом управление выполнением вывода нейросетей полностью передаётся вычислительному слою. Processor выполняет работу над данными, размещёнными в общей памяти, и возвращает в логический слой результаты обработки в виде структурированных метаданных.

При подтверждении факта ДТП логический слой формирует сообщение и передаёт его в Sender. Параллельно инициируется формирование видеоклипа фиксированной длительности на основе кадров скользящего окна и его сохранение во внешнем объектном хранилище. Операция сохранения видеоклипа выполняется асинхронно и не влияет на процесс отправки события. Освобождение областей общей памяти также осуществляется по принципу скользящего окна обработки видеопотока и не требует явного участия логического слоя, исключая утечки памяти при длительной работе системы.

4.4. Информационная модель АИС

В АИС используется информационная модель, определяющая состав данных, используемых на различных этапах обработки видеопотока, их логическую структуру и взаимосвязи, а также особенности жизненного цикла информационных объектов. Она ориентирована на in-memory информационные объекты, формируемые и используемые в процессе обработки видеопотоков, без применения персистентного хранения промежуточных данных обработки видеопотока. Исключение составляют видеоклипы, формируемые при подтверждении факта ДТП, которые являются производными артефактами событий и сохраняются во внешнем объектном хранилище. Состав информационных объектов, формируемых в процессе обработки видеопотока, а также их взаимосвязи представлены на рисунке 4.4.

Конфигурационные данные видеопотоков хранятся в CSV-файле и используются для инициализации и управления процессом захвата видеопотоков слоем Decoder. Файл включает следующие параметры:

- camera_id – уникальный идентификатор источника видеоданных;
- rtsp_url – адрес и параметры подключения к RTSP-потoku;
- camera_location – логическая привязка камеры к участку дорожной сети;
- enabled – текущее состояние соединения.

Параметры подключения к источникам видеонаблюдения задаются в конфигурационном CSV-файле, данные которого используются слоем Decoder при инициализации и последующем управлении видеопотоками. Для каждой камеры в конфигурации указывается уникальный идентификатор camera_id, адрес RTSP-потока и параметры подключения rtsp_url, а также camera_location, определяющий привязку источника к соответствующему участку дорожной сети. Текущее состояние подключения задается параметром enabled, на основании которого определяется необходимость обработки видеопотока конкретной камеры.

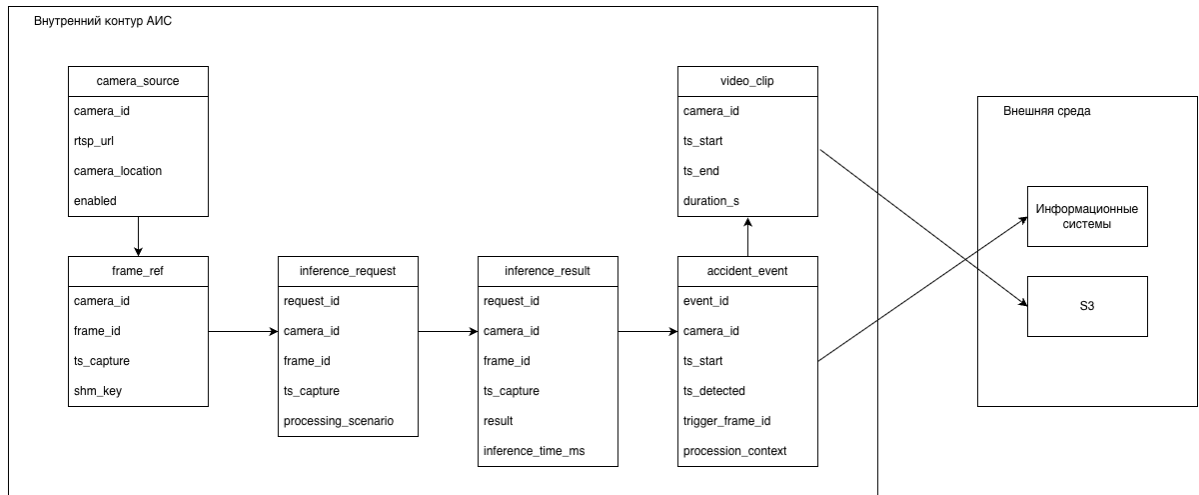


Рисунок 4.4 – Информационная модель автоматизированной информационной системы

В слое Logic идентификатор `camera_id` используется для связывания поступающих данных с состоянием конкретного видеопотока. Сведения `camera_location` обеспечивают его пространственную привязку и служат для получения метеорологических данных, соответствующих местоположению камеры.

Передача декодированных кадров между компонентами АИС осуществляется посредством информационного объекта `frame_ref`. Каждый такой объект связан с определенным источником видеонаблюдения через `camera_id`, содержит порядковый идентификатор кадра `frame_id` и временную метку его получения `ts_capture`. Сами пиксельные данные в объект не включаются: их размещение в общей памяти определяется атрибутом `shm_key`.

При обработке видеопотока Logic использует `camera_id`, `frame_id` и `ts_capture` для сохранения последовательности кадров и обновления состояния соответствующей камеры. Ссылка `shm_key` включается в запрос только в тех случаях, когда кадр необходимо передать на нейросетевую обработку. Получив `frame_ref`, Processor обращается непосредственно к указанной области общей памяти, поэтому вычислительный слой не зависит от механизма захвата и декодирования исходного видеопотока.

Обмен данными между Logic и Processor при выполнении нейросетевого вывода осуществляется посредством информационного объекта `inference_request`. Он формируется для конкретного кадра и определяет операцию, которую необходимо выполнить вычислительному слою. Для сопоставления запроса с результатом асинхронной обработки используется `request_id`, тогда как `camera_id`, `frame_id` и `ts_capture` сохраняют связь с исходным кадром и соответствующим видеопотоком. Доступ к пиксельным данным осуществляется по `shm_key`, указывающему на размещение кадра в общей памяти. Атрибут `processing_scenario` задает сценарий нейросетевой обработки и тем самым определяет набор вычислительных операций, выполняемых Processor.

После выполнения заданного сценария Processor формирует информационный объект `inference_result` и асинхронно возвращает его в Logic. Связь результата с исходным запросом сохраняется посредством `request_id`, а атрибуты `camera_id`, `frame_id` и `ts_capture` определяют принадлежность полученных данных конкретному кадру и видеопотоку. В `processing_scenario` фиксируется фактически выполненный сценарий обработки, поле `result` содержит сведения об обнаруженных объектах дорожной сцены и соответствующие показатели достоверности, а `inference_time_ms` характеризует продолжительность нейросетевого вывода.

Поступивший `inference_result` используется Logic для обновления состояния соответствующего видеопотока. Результаты отдельных кадров включаются во временные последовательности, на основе которых накапливаются признаки, предусмотренные разработанным во второй главе методом обнаружения ДТП. Тем самым результат нейросетевого вывода рассматривается не как самостоятельное решение о наличии происшествия, а как входная информация для последующего анализа динамики дорожной сцены.

После подтверждения ДТП в Logic создается информационный объект `accident_event`, предназначенный для передачи сведений о зарегистрированном событии во внешние информационные системы. Каждому событию

присваивается уникальный `event_id`, а его связь с источником видеоданных определяется по `camera_id`. Временные границы происшествия задаются атрибутами `ts_start` и `ts_detected`, при этом `trigger_frame_id` указывает кадр, на котором были выполнены условия подтверждения ДТП. Дополнительные сведения о процессе обнаружения сохраняются в `processing_context`, где фиксируются использованный сценарий обработки и нейросетевой метод.

Временные границы обнаруженного события определяются в Logic по содержимому скользящего окна, в пределах которого анализируется развитие дорожной ситуации. Значение `ts_detected` соответствует временной метке кадра, на котором были выполнены условия подтверждения ДТП. Начало интервала, использованного для накопления признаков события, фиксируется в `ts_start`.

Одновременно с формированием события Logic создает объект `video_clip`, содержащий связанный с ДТП фрагмент видеозаписи. Его временные границы задаются атрибутами `ts_start` и `ts_end`, продолжительность сохраняется в `duration_s`, а принадлежность определенному источнику видеонаблюдения устанавливается по `camera_id`. Видеоклип формируется из кадров, накопленных в соответствующем временном интервале, после чего сохраняется в объектном хранилище для последующего просмотра и анализа.

Объекты `accident_event` и `video_clip` имеют различное назначение и далее обрабатываются независимо: `accident_event` передается в слой внешнего взаимодействия и содержит только структурированные сведения о выявленном происшествии, без кадров и других объемных видеоданных. `video_clip`, напротив, остается в объектном хранилище и не передается через интерфейсы обмена событиями. Связь между ними устанавливается по `camera_id` и временным характеристикам, что позволяет сопоставить зарегистрированное ДТП с соответствующим фрагментом видеозаписи без включения самого видеоконтента в информационный объект события

Жизненный цикл информационных объектов в АИС определяется потоковым характером обработки видеоданных. Конфигурационные данные сохраняются в течение всего периода эксплуатации системы. В то же время

данные кадров и результаты обработки нейронной сетью создаются, используются и удаляются в рамках ограниченного временного интервала, необходимого для анализа видеопотока и принятия решений.

Информационная модель системы разработана с учетом потоковой обработки видеоданных и использует кратковременные информационные объекты, отражающие текущую дорожную обстановку. Отказ от постоянного хранения промежуточных данных позволяет сократить задержки при обработке и упростить архитектуру системы. Риск потери промежуточных данных вследствие сбоя компонентов минимизируется благодаря непрерывному характеру анализа видеопотока, а также использованию механизмов мониторинга процессов и трассировки.

4.5. Апробация и оценка эффективности разработанной АИС

Для подтверждения работоспособности разработанной автоматизированной информационной системы проведена экспериментальная апробация, направленная на оценку ее функциональной эффективности и производительности при увеличении числа одновременно обслуживаемых источников видеоданных.

Функциональная эффективность АИС. Поскольку в предыдущих главах диссертационного исследования разработанные методы восстановления изображения и обнаружения ДТП тестировались отдельно друг от друга необходимо провести оценку эффективности полного конвейера обработки видеоданных. Для моделирования ухудшенных условий наблюдения на исходные видеоданные накладывались синтетические дождевые артефакты. Их параметры соответствовали набору данных Rain800, использованному в третьей главе при оценке разработанного метода восстановления изображений.

Работа полного конвейера оценивалась в трех режимах: на исходных видеоданных без дополнительной обработки, на видеоданных с синтетическими дождевыми искажениями без восстановления и на тех же искаженных видеоданных после предварительной обработки разработанным методом восстановления изображений. Результаты экспериментальной оценки представлены в таблице 4.5.1.

Таблица 4.5.1. Результаты экспериментальной оценки влияния метода предобработки на обнаружение ДТП на видео

Конфигурация системы	Precision	Recall	F1	Скорость обработки на GPU (макс.), мс
Метод обнаружения ДТП без применения метода восстановления	0.85	0.88	0.86	20
Метод обнаружения ДТП без применения метода восстановления на шумных данных	0.76	0.72	0.74	20
Метод обнаружения ДТП с применением метода восстановления	0.77	0.86	0.8	52

Анализ полученных данных показывает, что влияние атмосферных дождевых осадков приводит к ухудшению качества обнаружения ДТП по всем рассматриваемым метрикам точности. При обработке зашумленных видеоданных без применения разработанного метода восстановления изображения наблюдается снижение точности обнаружения по всем наблюдаемым метрикам, что свидетельствует о потере значительной части информативных визуальных признаков дорожной сцены.

Использование разработанного метода восстановления изображения позволяет существенно повысить эффективность последующего обнаружения дорожно-транспортных происшествий. Полнота обнаружения возрастает до 0.86, что практически соответствует результатам, полученным на исходных видеоданных без погодных искажений. Одновременно наблюдается увеличение

точности классификации до 0.77 и рост F1-меры до 0.8. Несмотря на то что показатели качества остаются несколько ниже результатов, полученных на исходных изображениях, а скорость обработки полного конвейера превышает скорость обработки изображения без модуля восстановления, проведенный эксперимент подтверждает, что разработанный метод восстановления изображения сохраняет значительную часть признакового пространства, и повышает эффективность обнаружения ДТП в затрудненных погодных условиях.

Полученные результаты свидетельствуют, что совместное применение разработанных методов восстановления изображения и обнаружения дорожно-транспортных происшествий обеспечивает устойчивое функционирование полного конвейера обработки видеоданных в условиях ухудшенной видимости и позволяет существенно снизить влияние атмосферных зашумлений на качество распознавания аварийных ситуаций.

Производительность АИС. Для проверки работы полного программного конвейера при увеличении числа источников входных данных проведена серия контрольных нагрузочных замеров. В состав сценария входил сквозной вычислительный контур АИС: получение RTSP-потока, аппаратное декодирование H.264, работа разработанных методов восстановления изображения и обнаружения дорожно-транспортных происшествий, и отправка сигнала во внешнюю информационную систему.

Генерация требуемого числа независимых источников выполнялась на отдельном сервере с помощью MediaMTX [109]. Видеозаписи публиковались как потоки RTSP с кодеком H.264 и частотой 20 кадров/с. Исследовались два варианта входного конвейера: аппаратное декодирование средствами GStreamer при исходном разрешении 1920×1080 и аппаратное декодирование на базе NVIDIA DeepStream при исходном разрешении 640×360. Конвейер GStreamer обрабатывает один видеопоток одним экземпляром процесса декодировщика, вследствие чего представляет более устойчивое к потере входных данных решение, а высокое разрешение ориентировано на обработку потоков высокого

разрешения. В свою очередь DeepStream предоставляет возможность батчевого декодирования до 20 видеопотоков одним экземпляром процесса декодировщика. Однако при потере данных с одной из камер процесс нарушает свою работу, в связи с этим такое решение обеспечивает меньшую устойчивость, а пониженное разрешение ориентированно на эффективную утилизацию ресурсов. Конфигурация демо-стенда описана в таблице 4.5.2.

Таблица 4.5.2. Конфигурация и условия нагрузочных замеров

Элемент	Характеристика
Сервер АИС	Enzo RS720A-E11-RS12, форм-фактор 2U
Процессор	2 × AMD EPYC 7763, всего 128 ядер и 256 потоков
Графический ускоритель	NVIDIA H100, 80 ГБ HBM2e, PCIe Gen5 ×16
Оперативная память	512 ГБ DDR4-3200 ECC
Сетевые интерфейсы	2 × 10GbE
Параметры источников	RTSP, H.264, 20 кадр/с

Таблица 4.5.3. Результаты нагрузочных замеров двух конфигураций

Декодировщик	Число камер	Средняя загрузка CPU, %	Средняя загрузка GPU, %	Занято RAM, Гб	Занято VRAM, Гб
GStreamer	1	5	3	16,5	5,2
	10	32	37	24,9	21,0
	20	58	59	32,6	35,0
	30	77	77	38,6	45,0
	35	86	79	42,9	58,0
DeepStream	10	11	11	18,2	4,2
	20	21	20	21,0	7,5
	30	26	37	22,4	8,2

	40	32	48	11,0	24,7
	50	43	58	15,0	27,9

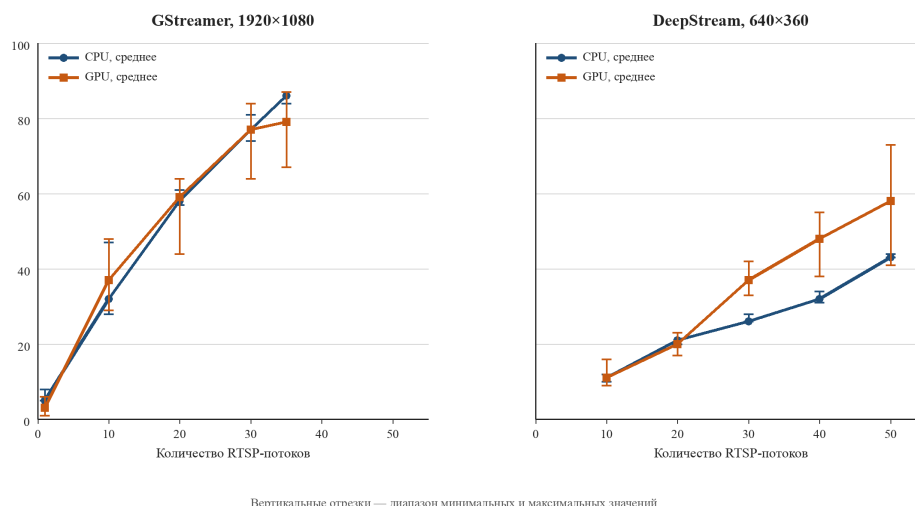


Рисунок 4.5 – Зависимость загрузки CPU и GPU от числа RTSP-потоков

Анализ конвейера GStreamer. При масштабировании системы с одной до 35 камер средняя загрузка центрального процессора возрастает с 5% до 86%, загрузка графического процессора – с 3% до 79%, а объем используемой видеопамяти – с 5,2 ГБ до 58 ГБ. При работе с 35 камерами средняя загрузка CPU достигает 86% при пиковом значении 87%, а средняя загрузка GPU составляет 79% при пиковом значении 87%. Данная конфигурация в рамках проведенной серии тестов показала основной ограничивающий ресурс центральный процессор.

Анализ конвейера DeepStream. При масштабировании системы с 10 до 50 средняя загрузка центрального процессора возрастает с 11% до 43%, графического процессора – с 11% до 58%, а объем используемой видеопамяти – с 4,2 ГБ до 27,9 ГБ. На максимальном из протестированных уровней пиковая загрузка центрального и графического процессоров достигает 44% и 73% соответственно. Текущая конфигурация продемонстрировала отсутствие насыщения аппаратных ресурсов при обработке 50 одновременных потоков.

Сопоставление вариантов. Конфигурация на базе GStreamer обеспечила обработку до 35 потоков, что характеризует верхнюю контрольную точку

испытания, но не рекомендуемую эксплуатационную нагрузку с гарантированным резервом. Конфигурация на базе DeepStream не продемонстрировала аппаратное насыщение при обработке 50 видеопотоков.

Оценка временной устойчивости. Для оценки временных характеристик АИС используется показатель сквозной задержки end-to-end, определяемый от момента передачи принятого видеокadra в аппаратный декодировщик до момента отправки сообщения о выявленном ДТП во внешнюю информационную систему.

Для оценки вычислительных затрат АИС были измерены временные характеристики выполнения разработанных методов восстановления изображений и обнаружения ДТП. Измерения проводились без учета времени ожидания в очередях, декодирования видеопотока и передачи выходных сообщений. Предварительная обработка кадра занимала не более 18 мс, выполнение нейросетевого вывода полного конвейера разработанных методов восстановления изображения и обнаружения ДТП – не более 52 мс, постобработка полученных результатов – не более 2 мс. Суммарное время выполнения обработки изображения не превышает 72 мс. Отдельно оценивались вычислительные затраты, связанные с функционированием АИС вне основного вычислительного блока. Аппаратное декодирование видеопотока H.264, передача кадра между компонентами системы и формирование выходного сообщения в совокупности занимали не более 5 мс. С учетом этих операций полное время обработки кадра АИС без учета ожидания в очереди не превышает 77 мс. Для конфигурации GStreamer результаты представлены в таблице 4.5.4.

При декодировании 20 потоков значение перцентиля задержки P95 не превышает 86 мс. При 30 потоках P95 возрастает приблизительно до 97 мс, а P99 – до 110 мс. Следовательно, конфигурация приближается к границе устойчивого режима: основная часть событий еще может обрабатываться в пределах установленного времени, однако редкие задержки уже превышают 100 мс.

Таблица 4.5.4. Роста латентности с использованием GStreamer при росте нагрузки

Число потоков	P50, мс	P95, мс	P99, мс
1	77	77	78
10	81	81	83
20	86	86	91
30	97	97	110
35	114	114	138

При 35 потоках средняя загрузка CPU достигает 86%, вследствие чего P95-латентность возрастает до 114 мс, а P99 – до 138 мс. Таким образом, максимальная исследованная ступень GStreamer является пограничной как по загрузке вычислительных ресурсов, так и по временным характеристикам. Для выполнения условия $P95 \leq 100$ мс эксплуатационную нагрузку данной конфигурации целесообразно ограничивать уровнем, не превышающим 30 потоков, либо предусматривать дополнительное масштабирование вычислительных компонентов.

Для конфигурации DeepStream результаты представлены в таблице 4.5.5.

Таблица 4.5.5. Рост латентности с использованием DeepStream при росте нагрузки

Число потоков	P50, мс	P95, мс	P99, мс
10	77	78	78
20	78	79	80
30	78	81	83
40	79	83	86
50	80	85	91

Для DeepStream увеличение числа потоков от 10 до 50 сопровождается ростом P95 с 78 до 85 мс, а P99 – с 78 до 91 мс. На максимальной исследованной ступени средняя загрузка наиболее занятого ресурса составляет 58%, поэтому нелинейного накопления очереди не ожидается. Временной резерв относительно ограничения $P95 \leq 100$ мс составляет около 15 мс.

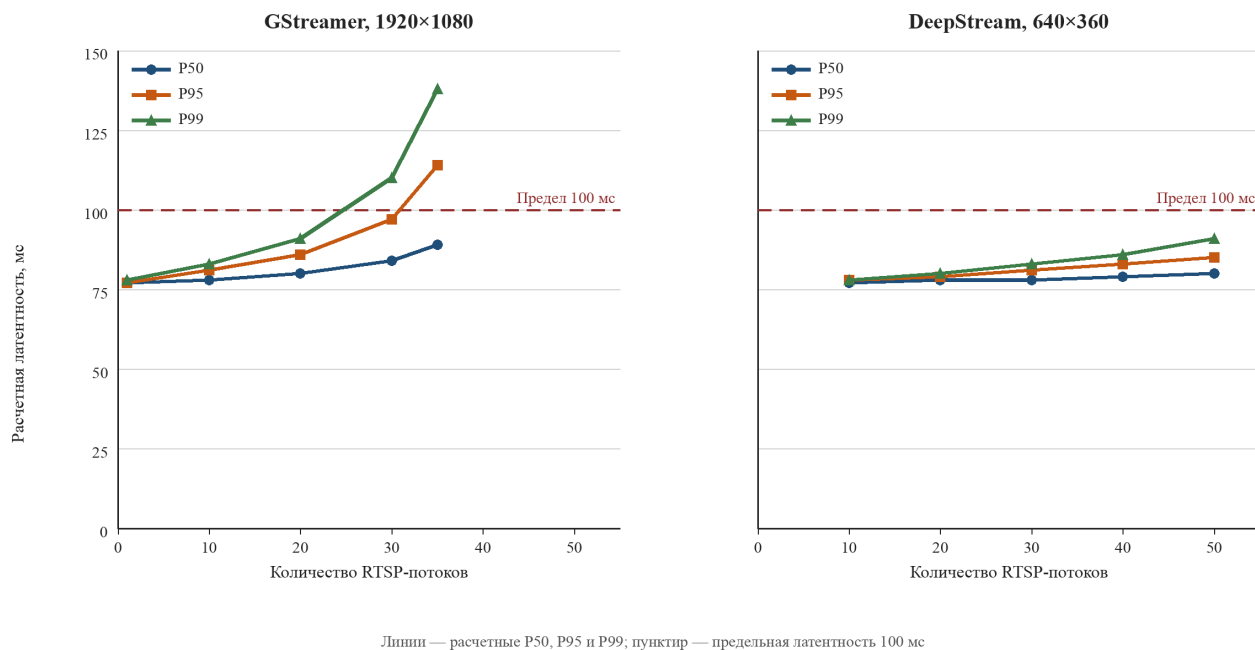


Рисунок 4.6 – Зависимость латентности от числа RTSP

На рисунке 4.6 отображено сравнение динамики роста латентности обеих конфигураций. Полученные зависимости показывают различный характер временного масштабирования двух потенциальных конфигураций АИС. Конфигурация GStreamer 1920×1080 сохраняет P95-латентность в пределах 100 мс приблизительно до 30 потоков, после чего рост загрузки CPU приводит к ускоренному увеличению времени ожидания. Конфигурация DeepStream 640×360 сохраняет ограничение $P95 \leq 100$ мс во всем исследованном диапазоне до 50 потоков.

4.6. Выводы по главе

В четвертой главе разработана архитектура автоматизированной информационной системы обнаружения ДТП условиях затрудненной видимости, обеспечивающая совместное применение разработанных методов обнаружения ДТП и восстановления изображений в условиях затруднённой видимости.

Для реализации разработанной архитектуры определены функциональная и информационная модели АИС. Масштабируемое взаимодействие компонентов обеспечивается асинхронной передачей запросов на нейросетевую обработку, хранением кадров в общей памяти без их повторного копирования. Пакетная обработка изображений позволяет одному экземпляру нейросетевого детектора принимать кадры с определенного количества видеопотоков и автоматически добавлять новые экземпляры детекторов при увеличении числа видеопотоков. Реализовано адаптивное формирование последовательности операций на основе сведений о погодных условиях, позволяющее выполнять ресурсоемкую восстановление только при ухудшении условий наблюдения.

Экспериментальная апробация полного конвейера обработки видеопотока показала, что при наложении дождевых искажений значение метрики Recall разработанного метода обнаружения ДТП снижается с 0,88 до 0,72. Применение разработанного метода восстановления повышает метрики Recall обнаружения ДТП на видеоданных с дождевыми искажениями с 0,72 до 0,86, Precision – с 0,76 до 0,77 и F1 – с 0,74 до 0,8. Вместе с тем время обработки кадра возрастает с 20 до 52 мс. Полученные результаты обосновывают адаптивную организацию обработки: применение метода восстановления позволяет в значительной степени восстановить качество обнаружения при атмосферных искажениях, но ее постоянное выполнение при нормальных условиях наблюдения приводит к избыточным вычислительным затратам.

Нагрузочные испытания полного вычислительного контура подтвердили возможность увеличения числа одновременно обрабатываемых видеопотоков без изменения структуры взаимодействия компонентов. При разрешении 1920×1080 и обработке до 30 видеопотоков, перцентиль задержки P95 составил 97 мс. При разрешении 640×360 АИС обрабатывала 50 видеопотоков при P95 = 85 мс и P99 = 91 мс. Полученные результаты определяют подтвержденные границы производительности исследованных конфигураций исследованных инструментов декодирования и доказывают возможность масштабирования АИС по числу источников видеоданных посредством изменения состава и количества вычислительных компонентов.

Таким образом, разработанная архитектура АИС объединяет методы восстановления изображений и обнаружения ДТП в единый процесс обработки видеопотока и обеспечивает последовательную обработку данных с адаптивным включением восстановления изображения при изменении условий наблюдения и масштабируемое взаимодействие функциональных компонентов в зависимости от количества видеопотоков. Полученные результаты подтверждают возможность совместного применения разработанных методов в задачах интеллектуального мониторинга дорожной обстановки.

Заключение

В рамках данного исследования была разработана и реализована автоматизированная информационная система устойчивого обнаружения дорожно-транспортных происшествий в условиях затрудненной видимости.

Ключевые достижения исследования включают:

1. Проведён анализ существующих методов обнаружения дорожно-транспортных происшествий и восстановления изображений при неблагоприятных погодных условиях. Установлено влияние атмосферных деградаций на снижение устойчивости алгоритмов компьютерного зрения и обоснована необходимость использования пространственно-временного анализа дорожной сцены и предварительной обработки видеоданных. (Задача № 1)
2. Разработан нейросетевой метод обнаружения дорожно-транспортных происшествий посредством анализа видеопотока с уличных городских камер. На первом этапе выполняются анализ динамики транспортных средств и формирование событий-кандидатов, на втором – нейросетевая классификация. Такая организация позволяет учитывать динамику развития дорожной сцены при анализе видеопоследовательности. Экспериментальная оценка на независимой тестовой выборке показала полноту выявления дорожно-транспортных происшествий не ниже 0,88 при времени обработки изображения не превышающим 20 мс. (Задача № 2, Положение № 1)
3. Разработан нейросетевой метод восстановления изображения от зашумлений, вызванных атмосферными осадками. В результате получено значение структурного сходства восстановленного изображения 0,78. После применения разработанного метода среднее значение mAP детекции объектов увеличилось с 0,20 на изображениях с дождевыми зашумлениями до 0,35. Время обработки изображения разработанным методом

восстановления не превышает 29 мс, что соответствует двукратному сокращению по сравнению с базовым методом. Полученные результаты подтверждают сохранение информативности визуальных признаков для последующего обнаружения объектов дорожной сцены. (Задача № 3, Положение № 2)

4. Разработана архитектура автоматизированной информационной системы обнаружения дорожно-транспортных происшествий в условиях затруднённой видимости. Система обеспечивает совместное применение разработанных методов, адаптивное включение восстановления изображения в зависимости от условий наблюдения и масштабируемое взаимодействие функциональных компонентов при изменении количества входных видеопотоков. Использование разработанного метода восстановления изображения в полном конвейере обработки данных повысило метрики Recall обнаружения ДТП на видеоданных с дождевыми искажениями с 0,72 до 0,86, Precision – с 0,76 до 0,77 и F1 – с 0,74 до 0,8. В исследованной конфигурации с разрешением входного видеопотока 640×360 обеспечена обработка 50 видеопотоков при $P95 = 85$ мс и $P99 = 91$ мс. (Задача № 4, Положение № 3)

Полученные результаты подтверждают возможность применения разработанных методов и архитектуры автоматизированной информационной системы в составе интеллектуальных транспортных систем. Совместное использование двухэтапного метода обнаружения ДТП, метода восстановления изображений и адаптивной организации вычислительного процесса обеспечивает обнаружение аварийных событий позволяет организовать обработку видеоданных с учетом изменения условий наблюдения и интенсивности поступающей нагрузки.

Список литературы

1. Дорожно-транспортная аварийность в Российской Федерации в 2025 году : информационно-аналитический обзор / К. С. Баканов, П. В. Ляхов, М. М. Исаев [и др.] ; под общ. ред. О. М. Портashникова, Д. В. Митрошина. – Москва : ФКУ «НЦ БДД МВД России», 2026. – 118 с. – ISBN 978-5-6053882-0-3. – URL: <https://media.mvd.ru/files/embed/19415255> (дата обращения: 15.03.2026).
2. Дорожно-транспортная аварийность в Российской Федерации в 2024 году : информационно-аналитический обзор / К. С. Баканов, П. В. Ляхов, А. С. Айсанов [и др.] ; под общ. ред. О. М. Портashникова, Д. В. Митрошина. – Москва : ФКУ «НЦ БДД МВД России», 2025. – 148 с. – ISBN 978-5-6046954-5-6. – URL: <https://media.mvd.ru/files/embed/11400230> (дата обращения: 15.03.2026).
3. Xu, R. Edge Video Analytics: A Survey on Applications, Systems and Enabling Techniques / R. Xu, S. Razavi, R. Zheng // IEEE Communications Surveys & Tutorials. – 2023. – Vol. 25. – No. 4. – P. 2951–2982. – DOI 10.1109/comst.2023.3323091.
4. Accident Prediction and Crash Recovery by using Car Black Box / P. S. Keerthi, S. A. Parveen, P. A. S. S. Sowmya [и др.] // International Journal of Innovative Technology and Exploring Engineering. – 2020. – Vol. 9. – No. 6. – P. 1394–1397. – DOI 10.35940/ijitee.f4215.049620.
5. Big Data and Deep Learning in Smart Cities: A Comprehensive Dataset for AI-Driven Traffic Accident Detection and Computer Vision Systems / V. Adewopo, N. Elsayed, Z. Elsayed [и др.] // 2024 IEEE 3rd International Conference on Computing and Machine Intelligence (ICMI). – 2024. – P. 1–6. – DOI 10.1109/icmi60790.2024.10586073.

6. *Lipatova, A. Researching Effective Systems and Methods for Detecting Drowsiness / A. Lipatova, M. Moseva, K. Kharrasov [и др.] // 2024 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF). – 2024. – P. 1–5. – DOI 10.1109/weconf61770.2024.10564573.*
7. Нейросетевая система определения сонливости человека по выражению лица / Б. Курбанов, А. С. Катасёв, Д. В. Катасёва, Б. Р. Зиннуров // Вестник технологического университета. – 2024. – Т. 27. – № 12. – С. 158–164. – DOI 10.55421/1998-7072_2024_27_12_158.
8. *S. M. K. M. Siam, Real-time accident detection and physiological signal monitoring to enhance motorbike safety and emergency response / K. I. Sumaiya, M. R. Al-Amin [и др.] // arXiv preprint. – 2024. – arXiv:2403.19085. – DOI 10.48550/arXiv.2403.19085. – URL: <https://arxiv.org/abs/2403.19085>*
9. Цяо, Ц. Метод интеграции мультимодальной информации для обнаружения и прогнозирования дорожно-транспортных происшествий на основе нейросетевого подхода / Ц. Цяо, И. С. Кружилов, О. Ю. Шамаева // Вестник МЭИ. – 2025. – № 3. – С. 153–160. – DOI 10.24160/1993-6982-2025-3-153-160.
10. Rahman, Z. A Real-Time Wrong-Way Vehicle Detection Based on YOLO and Centroid Tracking / Z. Rahman, A. M. Ami, M. A. Ullah // 2020 IEEE Region 10 Symposium (TENSYP). – 2020. – P. 916–920. – DOI 10.1109/tensymp50017.2020.9230463.
11. Ghahremannezhad, H. Real-Time Accident Detection in Traffic Surveillance Using Deep Learning / H. Ghahremannezhad, H. Shi, C. Liu // 2022 IEEE International Conference on Imaging Systems and Techniques (IST). – 2022. – P. 1–6. – DOI 10.1109/ist55454.2022.9827736.
12. Unsupervised Traffic Accident Detection in First-Person Videos / Y. Yao, M. Xu, Y. Wang [и др.] // 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). – 2019. – P. 273–280. – DOI 10.1109/iros40897.2019.8967556.

13. Vision Transformer-Based Multi-Phase Accident Detection and False Positive Mitigation / U. Jartarghar, D. Sanghvi, M. Nidgundi [и др.] // Research Square. – 2024. – DOI 10.21203/rs.3.rs-3903862/v1. – URL: <https://doi.org/10.21203/rs.3.rs-3903862/v1>
14. *Robles-Serrano, S.* Automatic Detection of Traffic Accidents from Video Using Deep Learning Techniques / S. Robles-Serrano, G. Sanchez-Torres, J. Branch-Bedoya // Computers. – 2021. – Vol. 10. – No. 11. – Art. 148. – DOI 10.3390/computers10110148.
15. Deep Learning applied to Road Accident Detection with Transfer Learning and Synthetic Images / T. Tamagusko, M. G. Correia, M. A. Huynh [и др.] // Transportation Research Procedia. – 2022. – Vol. 64. – P. 90–97. – DOI 10.1016/j.trpro.2022.09.012.
16. *Tan, M.* EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks / M. Tan, Q. Le // Proceedings of the 36th International Conference on Machine Learning. – 2019. – Vol. 97. – P. 6105–6114. – URL: <https://proceedings.mlr.press/v97/tan19a.html>
17. MobileNetV2: Inverted Residuals and Linear Bottlenecks / M. Sandler, A. Howard, M. Zhu [и др.] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2018. – P. 4510–4520. – DOI 10.1109/cvpr.2018.00474.
18. Real-Time Event-Driven Road Traffic Monitoring System Using CCTV Video Analytics / M. Tahir, Y. Qiao, N. Kanwal [и др.] // IEEE Access. – 2023. – Vol. 11. – P. 139097–139111. – DOI 10.1109/access.2023.3340144.
19. *Pawar, K.* Deep learning based detection and localization of road accidents from traffic surveillance videos / K. Pawar, V. Attar // ICT Express. – 2022. – Vol. 8. – No. 3. – P. 379–387. – DOI 10.1016/j.icte.2021.11.004.
20. *Singh, D.* Deep Spatio-Temporal Representation for Detection of Road Accidents Using Stacked Autoencoder / D. Singh, C. K. Mohan // IEEE Transactions on Intelligent Transportation Systems. – 2019. – Vol. 20. – No. 3. – P. 879–887. – DOI 10.1109/tits.2018.2835308.

21. *Essam, M.* Detection of Road Traffic Crashes based on Collision Estimation / M. Essam, N. M. Ghanem, M. A. Ismail // Artificial Intelligence and Machine Learning. – 2022. – P. 169–179. – DOI 10.5121/csit.2022.121213.
22. *Moreno Pallares, M.* VARVO: a Novel Method for the Fast Detection of Vehicle Crash Events from Video Only Data / M. Moreno Pallares, S. G. Yoo, W. G. Aguilar // Revista Politécnica. – 2023. – Vol. 52. – No. 1. – P. 25–34. – DOI 10.33333/rp.vol52n1.03.
23. Real-Time Vehicle Accident Recognition from Traffic Video Surveillance using YOLOV8 and OpenCV / D. T. Mane, S. Sangve, S. Kandhare [и др.] // International Journal on Recent and Innovation Trends in Computing and Communication. – 2023. – Vol. 11. – No. 5s. – P. 250–258. – DOI 10.17762/ijritcc.v11i5s.6651.
24. *Zhang, Y.* MOTRv2: Bootstrapping End-to-End Multi-Object Tracking by Pretrained Object Detectors / Y. Zhang, T. Wang, X. Zhang // 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2023. – P. 22056–22065. – DOI 10.1109/cvpr52729.2023.02112.
25. Robust Object Detection in Challenging Weather Conditions / H. Gupta, O. Kotlyar, H. Andreasson [и др.] // 2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). – 2024. – P. 7508–7517. – DOI 10.1109/wacv57701.2024.00735.
26. *Ziyadinov, V.* Noise Immunity and Robustness Study of Image Recognition Using a Convolutional Neural Network / V. Ziyadinov, M. Tereshonok // Sensors. – 2022. – Vol. 22. – No. 3. – Art. 1241. – DOI 10.3390/s22031241.
27. *Ziyadinov, V.* Low-Pass Image Filtering to Achieve Adversarial Robustness / V. Ziyadinov, M. Tereshonok // Sensors. – 2023. – Vol. 23. – No. 22. – Art. 9032. – DOI 10.3390/s23229032.
28. A comprehensive survey on image dehazing for different atmospheric scattering models / S. An, X. Huang, L. Cao [и др.] // Multimedia Tools and Applications. – 2023. – Vol. 83. – No. 14. – P. 40963–40993. – DOI 10.1007/s11042-023-17292-8.

29. *Narasimhan, S. G.* Contrast restoration of weather degraded images / S. G. Narasimhan, S. K. Nayar // ACM SIGGRAPH ASIA 2008 courses on - SIGGRAPH Asia '08. – 2008. – P. 1–12. – DOI 10.1145/1508044.1508114.
30. *Hanafy, M. E.* Detailed effects of scattering and absorption by haze and aerosols in the atmosphere on the average point spread function of an imaging system / M. E. Hanafy, M. C. Roggemann, D. O. Guney // Journal of the Optical Society of America A. – 2014. – Vol. 31. – No. 6. – P. 1312. – DOI 10.1364/josaa.31.001312.
31. Sparse Sampling Transformer with Uncertainty-Driven Ranking for Unified Removal of Raindrops and Rain Streaks / S. Chen, T. Ye, J. Bai [и др.] // 2023 IEEE/CVF International Conference on Computer Vision (ICCV). – 2023. – P. 13060–13071. – DOI 10.1109/iccv51070.2023.01205.
32. Single Image Deraining: From Model-Based to Data-Driven and Beyond / W. Yang, R. T. Tan, S. Wang [и др.] // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2021. – Vol. 43. – No. 11. – P. 4059–4077. – DOI 10.1109/tpami.2020.2995190.
33. Rain Rendering for Evaluating and Improving Robustness to Bad Weather / M. Tremblay, S. S. Halder, R. de Charette [и др.] // International Journal of Computer Vision. – 2021. – Vol. 129. – No. 2. – P. 341–360. – DOI 10.1007/s11263-020-01366-3.
34. MSP-Former: Multi-Scale Projection Transformer for Single Image Desnowing / S. Chen, T. Ye, Y. Liu [и др.] // ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). – 2023. – P. 1–5. – DOI 10.1109/icassp49357.2023.10095605.
35. Uncertainty-Driven Dynamic Degradation Perceiving and Background Modeling for Efficient Single Image Desnowing / S. Chen, T. Ye, C. Xue [и др.] // Proceedings of the 31st ACM International Conference on Multimedia. – 2023. – P. 4269–4280. – DOI 10.1145/3581783.3612003.

36. LiDAR Snowfall Simulation for Robust 3D Object Detection / M. Hahner, C. Sakaridis, M. Bijelic [и др.] // 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2022. – P. 16343–16353. – DOI 10.1109/cvpr52688.2022.01588.
37. Snow-CLOCs: Camera-LiDAR Object Candidate Fusion for 3D Object Detection in Snowy Conditions / X. Fan, D. Xiao, Q. Li [и др.] // Sensors. – 2024. – Vol. 24. – No. 13. – Art. 4158. – DOI 10.3390/s24134158.
38. CF-YOLO: Cross Fusion YOLO for Object Detection in Adverse Weather With a High-Quality Real Snow Dataset / Q. Ding, P. Li, X. Yan [и др.] // IEEE Transactions on Intelligent Transportation Systems. – 2023. – Vol. 24. – No. 10. – P. 10749–10759. – DOI 10.1109/tits.2023.3285035.
39. *Garg, K. Vision and Rain* / K. Garg, S. K. Nayar // International Journal of Computer Vision. – 2007. – Vol. 75. – No. 1. – P. 3–27. – DOI 10.1007/s11263-006-0028-6.
40. *Jose Valanarasu, J. M. TransWeather: Transformer-based Restoration of Images Degraded by Adverse Weather Conditions* / J. M. Jose Valanarasu, R. Yasarla, V. M. Patel // 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2022. – P. 2343–2353. – DOI 10.1109/cvpr52688.2022.00239.
41. Learning Multiple Adverse Weather Removal via Two-stage Knowledge Learning and Multi-contrastive Regularization: Toward a Unified Model / W. T. Chen, Z. K. Huang, C. C. Tsai [и др.] // 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2022. – P. 17632–17641. – DOI 10.1109/cvpr52688.2022.01713.
42. All-In-One Image Restoration for Unknown Corruption / B. Li, X. Liu, P. Hu [и др.] // 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2022. – P. 17431–17441. – DOI 10.1109/cvpr52688.2022.01693.
43. MAXIM: Multi-Axis MLP for Image Processing / Z. Tu, H. Talebi, H. Zhang [и др.] // 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2022. – P. 5759–5770. – DOI 10.1109/cvpr52688.2022.00568.

44. *Chen, C. P.* SensingGAN: A Lightweight GAN for Real-time Rain Removal with Self-attention / C. P. Chen // GitHub. – URL: <https://github.com/chenabei0421/SensingGAN> (дата обращения: 28.12.2024).
45. Image-Adaptive YOLO for Object Detection in Adverse Weather Conditions / W. Liu, G. Ren, R. Yu [и др.] // Proceedings of the AAAI Conference on Artificial Intelligence. – 2022. – Vol. 36. – No. 2. – P. 1792–1800. – DOI 10.1609/aaai.v36i2.20072.
46. GDIP: Gated Differentiable Image Processing for Object Detection in Adverse Conditions / S. Kalwar, D. Patel, A. Aanegola [и др.] // 2023 IEEE International Conference on Robotics and Automation (ICRA). – 2023. – P. 7083–7089. – DOI 10.1109/icra48891.2023.10160356.
47. TogetherNet: Bridging Image Restoration and Object Detection Together via Dynamic Enhancement Learning / Y. Wang, X. Yan, K. Zhang [и др.] // Computer Graphics Forum. – 2022. – Vol. 41. – No. 7. – P. 465–476. – DOI 10.1111/cgf.14692.
48. A Survey on All-in-One Image Restoration: Taxonomy, Evaluation and Future Trends / J. Jiang, Z. Zuo, G. Wu [и др.] // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2025. – Vol. 47. – No. 12. – P. 11892–11911. – DOI 10.1109/tpami.2025.3598132.
49. Simple Baselines for Image Restoration / L. Chen, X. Chu, X. Zhang [и др.] // Computer Vision – ECCV 2022. – 2022. – P. 17–33. – DOI 10.1007/978-3-031-20071-7_2.
50. Revitalizing Convolutional Network for Image Restoration / Y. Cui, W. Ren, X. Cao [и др.] // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2024. – Vol. 46. – No. 12. – P. 9423–9438. – DOI 10.1109/tpami.2024.3419007.
51. Progressive Image Deraining Networks: A Better and Simpler Baseline / D. Ren, W. Zuo, Q. Hu [и др.] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2019. – P. 3932–3941. – DOI 10.1109/cvpr.2019.00406.

52. Multi-Stage Progressive Image Restoration / S. W. Zamir, A. Arora, S. Khan [и др.] // 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2021. – P. 14816–14826. – DOI 10.1109/cvpr46437.2021.01458.
53. From heavy rain removal to detail restoration: A faster and better network / Y. Wen, T. Gao, J. Zhang [и др.] // Pattern Recognition. – 2024. – Vol. 148. – Art. 110205. – DOI 10.1016/j.patcog.2023.110205.
54. AOD-Net: All-in-One Dehazing Network / B. Li, X. Peng, Z. Wang [и др.] // 2017 IEEE International Conference on Computer Vision (ICCV). – 2017. – P. 4780–4788. – DOI 10.1109/iccv.2017.511.
55. Vision Transformers for Single Image Dehazing / Y. Song, Z. He, H. Qian [и др.] // IEEE Transactions on Image Processing. – 2023. – Vol. 32. – P. 1927–1941. – DOI 10.1109/tip.2023.3256763.
56. Focal Network for Image Restoration / Y. Cui, W. Ren, X. Cao [и др.] // 2023 IEEE/CVF International Conference on Computer Vision (ICCV). – 2023. – P. 12955–12965. – DOI 10.1109/iccv51070.2023.01195.
57. TANet: Triplet Attention Network for All-In-One Adverse Weather Image Restoration / H. H. Wang, F. J. Tsai, Y. Y. Lin [и др.] // Computer Vision – ACCV 2024. – 2025. – P. 3–19. – DOI 10.1007/978-981-96-0911-6_1.
58. *Haouassi, S.* An Efficient Attentional Image Dehazing Deep Network Using Two Color Space (ADMC2-net) / S. Haouassi, D. Wu // Sensors. – 2024. – Vol. 24. – No. 2. – Art. 687. – DOI 10.3390/s24020687.
59. *Wang, Y.* CasDyF-Net: Image Dehazing via Cascaded Dynamic Filters / Y. Wang, B. He // 2024 IEEE 8th International Conference on Vision, Image and Signal Processing (ICVISIP). – 2024. – P. 1–8. – DOI 10.1109/icvisip64524.2024.10959384.
60. MixDehazeNet: Mix Structure Block For Image Dehazing Network / L. Lu, Q. Xiong, B. Xu [и др.] // 2024 International Joint Conference on Neural Networks (IJCNN). – 2024. – P. 1–10. – DOI 10.1109/ijcnn60899.2024.10651326.

61. Generative Adversarial Nets / I. J. Goodfellow, J. Pouget-Abadie, M. Mirza [и др.] // *Advances in Neural Information Processing Systems*. – 2014. – Vol. 27. – P. 2672–2680. – URL: https://proceedings.neurips.cc/paper_files/paper/2014/hash/f033ed80deb0234979a61f95710dbe25-Abstract.html
62. *Zhang, H.* Image De-Raining Using a Conditional Generative Adversarial Network / H. Zhang, V. Sindagi, V. M. Patel // *IEEE Transactions on Circuits and Systems for Video Technology*. – 2020. – Vol. 30. – No. 11. – P. 3943–3956. – DOI 10.1109/tcsvt.2019.2920407.
63. Attentive Generative Adversarial Network for Raindrop Removal from A Single Image / R. Qian, R. T. Tan, W. Yang [и др.] // *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. – 2018. – P. 2482–2491. – DOI 10.1109/cvpr.2018.00263.
64. DerainCycleGAN: Rain Attentive CycleGAN for Single Image Deraining and Rainmaking / Y. Wei, Z. Zhang, Y. Wang [и др.] // *IEEE Transactions on Image Processing*. – 2021. – Vol. 30. – P. 4788–4801. – DOI 10.1109/tip.2021.3074804.
65. Multi-weather Image Restoration via Domain Translation / P. W. Patil, S. Gupta, S. Rana [и др.] // *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*. – 2023. – P. 21639–21648. – DOI 10.1109/iccv51070.2023.01983.
66. Restormer: Efficient Transformer for High-Resolution Image Restoration / S. W. Zamir, A. Arora, S. Khan [и др.] // *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. – 2022. – P. 5718–5729. – DOI 10.1109/cvpr52688.2022.00564.
67. Pre-Trained Image Processing Transformer / H. Chen, Y. Wang, T. Guo [и др.] // *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. – 2021. – P. 12294–12305. – DOI 10.1109/cvpr46437.2021.01212.
68. Instruct-IPT: All-in-One Image Processing Transformer via Weight Modulation / Y. Tian, J. Han, H. Chen [и др.] // *arXiv preprint*. – 2024. – arXiv:2407.00676. – DOI 10.48550/arXiv.2407.00676. – URL: <https://arxiv.org/abs/2407.00676>

69. Uformer: A General U-Shaped Transformer for Image Restoration / Z. Wang, X. Cun, J. Bao [и др.] // 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2022. – P. 17662–17672. – DOI 10.1109/cvpr52688.2022.01716.
70. LMQFormer: A Laplace-Prior-Guided Mask Query Transformer for Lightweight Snow Removal / J. Lin, N. Jiang, Z. Zhang [и др.] // IEEE Transactions on Circuits and Systems for Video Technology. – 2023. – Vol. 33. – No. 11. – P. 6225–6235. – DOI 10.1109/tcsvt.2023.3264824.
71. SnowFormer: Context Interaction Transformer with Scale-awareness for Single Image Desnowing / S. Chen, T. Ye, Y. Liu [и др.] // arXiv preprint. – 2022. – arXiv:2208.09703. – DOI 10.48550/arXiv.2208.09703. – URL: <https://arxiv.org/abs/2208.09703>
72. Joint multi-dimensional dynamic attention and transformer for general image restoration / H. Zhang, X. Zhang, N. Cai [и др.] // Computer Vision and Image Understanding. – 2025. – Vol. 261. – Art. 104491. – DOI 10.1016/j.cviu.2025.104491.
73. PromptIR: Prompting for All-in-One Image Restoration / V. Potlapalli, S. W. Zamir, S. Khan [и др.] // Advances in Neural Information Processing Systems. – 2023. – Vol. 36. – P. 71275–71293. – DOI 10.52202/075280-3121.
74. *Conde, M. V.* InstructIR: High-Quality Image Restoration Following Human Instructions / M. V. Conde, G. Geigle, R. Timofte // Computer Vision – ECCV 2024. – 2025. – P. 1–21. – DOI 10.1007/978-3-031-72764-1_1.
75. Denoising Diffusion Restoration Models / B. Kawar, M. Elad, S. Ermon [и др.] // Advances in Neural Information Processing Systems. – 2022. – Vol. 35. – P. 23593–23606. – DOI 10.52202/068431-1714.
76. Residual Denoising Diffusion Models / J. Liu, Q. Wang, H. Fan [и др.] // 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2024. – P. 2773–2783. – DOI 10.1109/cvpr52733.2024.00268.

77. DiffBIR: Toward Blind Image Restoration with Generative Diffusion Prior / X. Lin, J. He, Z. Chen [и др.] // Computer Vision – ECCV 2024. – 2025. – P. 430–448. – DOI 10.1007/978-3-031-73202-7_25.
78. Rethinking Real-world Image Deraining via An Unpaired Degradation-Conditioned Diffusion Model / Y. Shen, M. Wei, Y. Wang [и др.] // arXiv preprint. – 2023. – arXiv:2301.09430. – DOI 10.48550/arXiv.2301.09430. – URL: <https://arxiv.org/abs/2301.09430>
79. All-in-One Weather-Degraded Image Restoration Via Adaptive Degradation-Aware Self-Prompting Model / Y. Wen, T. Gao, Z. Li [и др.] // IEEE Transactions on Multimedia. – 2025. – Vol. 27. – P. 3343–3355. – DOI 10.1109/tmm.2025.3535316.
80. *Özdenizci, O.* Restoring Vision in Adverse Weather Conditions With Patch-Based Denoising Diffusion Models / O. Özdenizci, R. Legenstein // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2023. – Vol. 45. – No. 8. – P. 10346–10357. – DOI 10.1109/tpami.2023.3238179.
81. DiffIR: Efficient Diffusion Model for Image Restoration / B. Xia, Y. Zhang, S. Wang [и др.] // 2023 IEEE/CVF International Conference on Computer Vision (ICCV). – 2023. – P. 13049–13059. – DOI 10.1109/iccv51070.2023.01204.
82. DPM-Solver: A Fast ODE Solver for Diffusion Probabilistic Model Sampling in Around 10 Steps / C. Lu, Y. Zhou, F. Bao [и др.] // Advances in Neural Information Processing Systems. – 2022. – Vol. 35. – P. 5775–5787. – DOI 10.52202/068431-0418.
83. Consistency Models / Y. Song, P. Dhariwal, M. Chen [и др.] // Proceedings of the 40th International Conference on Machine Learning. – 2023. – Vol. 202. – P. 32211–32252. – URL: <https://proceedings.mlr.press/v202/song23a.html> (дата обращения: 15.09.2026).
84. Swin Transformer V2: Scaling Up Capacity and Resolution / Z. Liu, H. Hu, Y. Lin [и др.] // 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2022. – P. 11999–12009. – DOI 10.1109/cvpr52688.2022.01170.

85. EfficientFormer: Vision Transformers at MobileNet Speed / Y. Li, G. Yuan, Y. Wen [и др.] // Advances in Neural Information Processing Systems. – 2022. – Vol. 35. – P. 12934–12949. – DOI 10.52202/068431-0940.
86. EfficientViT: Memory Efficient Vision Transformer with Cascaded Group Attention / X. Liu, H. Peng, N. Zheng [и др.] // 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2023. – P. 14420–14430. – DOI 10.1109/cvpr52729.2023.01386.
87. Traffic-Net // GitHub. – URL: <https://github.com/OlafenwaMoses/Traffic-Net> (дата обращения: 05.06.2024).
88. A system of vision sensor based deep neural networks for complex driving scene analysis in support of crash risk assessment and prevention / M. M. Karim, Y. Li, R. Qin [и др.] // arXiv preprint. – 2021. – arXiv:2106.10319. – DOI 10.48550/arXiv.2106.10319. – URL: <https://arxiv.org/abs/2106.10319>
89. Run, Don't Walk: Chasing Higher FLOPS for Faster Neural Networks / J. Chen, S. H. Kao, H. He [и др.] // 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2023. – P. 12021–12031. – DOI 10.1109/cvpr52729.2023.01157.
90. MobileOne: An Improved One millisecond Mobile Backbone / P. K. A. Vasu, J. Gabriel, J. Zhu [и др.] // 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2023. – P. 7907–7917. – DOI 10.1109/cvpr52729.2023.00764.
91. Rep ViT: Revisiting Mobile CNN From ViT Perspective / A. Wang, H. Chen, Z. Lin [и др.] // 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2024. – P. 15909–15920. – DOI 10.1109/cvpr52733.2024.01506.
92. ConvNeXt V2: Co-designing and Scaling ConvNets with Masked Autoencoders / S. Woo, S. Debnath, R. Hu [и др.] // 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2023. – P. 16133–16142. – DOI 10.1109/cvpr52729.2023.01548.

93. Masked Autoencoders Are Scalable Vision Learners / K. He, X. Chen, S. Xie [и др.] // 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2022. – P. 15979–15988. – DOI 10.1109/cvpr52688.2022.01553.
94. Image Restoration via Frequency Selection / Y. Cui, W. Ren, X. Cao [и др.] // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2024. – Vol. 46. – No. 2. – P. 1093–1108. – DOI 10.1109/tpami.2023.3330416.
95. Efficient Frequency Domain-based Transformers for High-Quality Image Deblurring / L. Kong, J. Dong, J. Ge [и др.] // 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2023. – P. 5886–5895. – DOI 10.1109/cvpr52729.2023.00570.
96. *Chi, L.* Fast Fourier Convolution / L. Chi, B. Jiang, Y. Mu // Advances in Neural Information Processing Systems. – 2020. – Vol. 33. – P. 4479–4488. – URL: <https://proceedings.neurips.cc/paper/2020/hash/2fd5d41ec6cfab47e32164d5624269b1-Abstract.html>
97. ByteTrack: Multi-object Tracking by Associating Every Detection Box / Y. Zhang, P. Sun, Y. Jiang [и др.] // Computer Vision – ECCV 2022. – 2022. – P. 1–21. – DOI 10.1007/978-3-031-20047-2_1.
98. Observation-Centric SORT: Rethinking SORT for Robust Multi-Object Tracking / J. Cao, J. Pang, X. Weng [и др.] // 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2023. – P. 9686–9696. – DOI 10.1109/cvpr52729.2023.00934.
99. DoTA: Unsupervised Detection of Traffic Anomaly in Driving Videos / Y. Yao, X. Wang, M. Xu [и др.] // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2023. – Vol. 45. – No. 1. – P. 444–459. – DOI 10.1109/tpami.2022.3150763.
100. Москва: среднемесячные климатические данные // Гидрометцентр России : официальный сайт. – URL: <https://meteoinfo.ru/categ-articles/98-climate-cat/1709-moskva-climat> (дата обращения: 25.02.2025).

101. A survey on knowledge distillation: Recent advancements / A. Moslemi, A. Briskina, Z. Dang [и др.] // Machine Learning with Applications. – 2024. – Vol. 18. – Art. 100605. – DOI 10.1016/j.mlwa.2024.100605.
102. Decoupled Knowledge Distillation / B. Zhao, Q. Cui, R. Song [и др.] // 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2022. – P. 11943–11952. – DOI 10.1109/cvpr52688.2022.01165.
103. PTQ4ViT: Post-training Quantization for Vision Transformers with Twin Uniform Quantization / Z. Yuan, C. Xue, Y. Chen [и др.] // Computer Vision – ECCV 2022. – 2022. – P. 191–207. – DOI 10.1007/978-3-031-19775-8_12.
104. FQ-ViT: Post-Training Quantization for Fully Quantized Vision Transformer / Y. Lin, T. Zhang, P. Sun [и др.] // Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence. – 2022. – P. 1173–1179. – DOI 10.24963/ijcai.2022/164.
105. Post-Training Quantization on Diffusion Models / Y. Shang, Z. Yuan, B. Xie [и др.] // 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2023. – P. 1972–1981. – DOI 10.1109/cvpr52729.2023.00196.
106. Q-Diffusion: Quantizing Diffusion Models / X. Li, Y. Liu, L. Lian [и др.] // 2023 IEEE/CVF International Conference on Computer Vision (ICCV). – 2023. – P. 17489–17499. – DOI 10.1109/iccv51070.2023.01608.
107. BDD100K: A Diverse Driving Dataset for Heterogeneous Multitask Learning / F. Yu, H. Chen, X. Wang [и др.] // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2020. – P. 2633–2642. – DOI 10.1109/cvpr42600.2020.00271.
108. An Edge-Side Real-Time Video Analytics System With Dual Computing Resource Control / C. Hu, R. Lu, Q. Sang [и др.] // IEEE Transactions on Computers. – 2023. – Vol. 72. – No. 12. – P. 3399–3415. – DOI 10.1109/tc.2023.3301136.
109. MediaMTX: ready-to-use media server and media proxy // GitHub. – URL: <https://github.com/bluenviron/mediamtx> (дата обращения: 22.01.2026).

Приложение А. Модель угроз возникновения дорожно-транспортных происшествий

Улицы										
Кейс	Тип происшеств ия	Кол- во учас тник ов	Пострад авшие	Поведение и оформление	Препятстви е движению	Продол жительн ость трафика нарушен ия	Зон а вли яни я	Каскад ный риск	Необходи мость вызова служб	Приорит ет для распозна вания
1. Столкновен ие с остановкой, без вызова ГИБДД	Наезд сзади, повреждени е бампера	2	Нет	Остановка, самостоятель ное оформление европротоко ла на обочине	Минимальн ое (часть правой полосы)	10–15 минут	Лок аль ная	Отсутс твует	Отсутств ует	Низкий / Средний

2. Блокировка одной полосы и вызов ГИБДД	Боковое столкновение при перестроении	2	Нет	Остановка с перекрытием полос движения, вызов служб	Частичное (1 из 2 полос занята)	30–40 минут	До 500 м	Средний	Отсутствует	Средний
3. Наезд на пешехода/инфраструктуру	Наезд на пешехода/повреждение инфраструктуры	1	Пешеход/водитель	Остановка с перекрытием полос движения, вызов служб	Частичное (остановка в полосе)	Около 1 часа	До 1 км	Средний	Требуется участие ГИБДД	Высокий
4. Цепное ДТП, полная блокировка	Цепная авария	3 и более авто	Возможны травмы	Остановка с перекрытием полос движения, вызов служб	Полная блокировка 2 полос	>1 часа	До 2–3 км	Высокий	Требуется участие всех экстренных служб	Высокий

5. Столкновение при левом повороте	Левый поворот, конфликт с прямым потоком	2	Нет	Остановка с перекрытием полос движения, вызов служб	Частичное	20 мин	До 300 м	Средний	Требуется участие ГИБДД	Средний
6. Наезд на пешехода на переходе	Наезд при разрешающем сигнале	1	Пешеход (травма)	Остановка с перекрытием полос движения, вызов служб	Нет	30 мин	Локальная	Низкий	Требуется участие ГИБДД и скорой	Высокий
7. Лобовое на реверсе	Выезд на реверс	2	Лёгкие травмы	Остановка с перекрытием полос движения, вызов служб	Полная	50 мин	До 1 км	Высокий	Требуется участие ГИБДД и скорой	Высокий

8. Наезд на велосипедиста	Выезд из двора	1	Велосипедист (лёгкая травма)	Остановка с перекрытием полос движения, вызов служб	Нет	20 мин	Локальная	Отсутствует	Требуется участие скорой	Средний
9. ДТП с фурой	Фура и 2 легковых	3	Госпитализация	Остановка с перекрытием полос движения, вызов служб	Полная	>1.5 ч	2–4 км	Очень высокой	Требуется участие всех экстренных служб	Высокий
10. Цепное столкновение	Плохая видимость	4	Возможны	Остановка с перекрытием полос движения, вызов служб	Частичное	~1 ч	До 2 км	Высокий	Требуется участие ГИБДД и скорой	Высокий

11. Наезд у остановки	Пассажир при посадке	1	Тяжёлая травма	Остановка с перекрытием полос движения, вызов служб	Нет	30 мин	Локаль ная	Низкий	Требуется участие ГИБДД и скорой	Высокий
12. ДТП в тоннеле	Столкновение	2	Нет	Остановка с перекрытием полос движения, вызов служб	Полная	>1 ч	3 км+	Высокий	Требуется участие ГИБДД и скорой	Высокий

Приложение Б. Свидетельства о государственной регистрации программ для ЭВМ

РОССИЙСКАЯ ФЕДЕРАЦИЯ



RU2025666153

ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ
ГОСУДАРСТВЕННАЯ РЕГИСТРАЦИЯ ПРОГРАММЫ ДЛЯ ЭВМ

Номер регистрации (свидетельства):
2025666153

Дата регистрации: **24.06.2025**

Номер и дата поступления заявки:
2025665050 04.06.2025

Дата публикации и номер бюллетеня:
24.06.2025 Бюл. № 7

Контактные реквизиты:
нет

Автор(ы):

Харрасов Камиль Раисович (RU)

Правообладатель(и):

**Ордена Трудового Красного Знамени
федеральное государственное бюджетное
образовательное учреждение высшего
образования «Московский технический
университет связи и информатики» (RU)**

Название программы для ЭВМ:

**Программное обеспечение для обнаружения аномальных событий на улично-дорожной сети
посредством обработки видеоконтента нейросетевым методом**

Реферат:

Современные системы определения дорожно-транспортных происшествий обладают рядом ограничений, включая недостаточную точность, задержки в обработке данных и сложности в выявлении аварийных ситуаций в реальном времени, что снижает эффективность мониторинга дорожной обстановки. Разработанная программа, использующая нейро-сетевой метод обнаружения ДТП, использует модифицированные операции сверки для ускоренной обработки визуальных данных, что позволяет повысить скорость анализа каждого кадра и способствует более оперативному выявлению аварийных ситуаций и реагированию на дорожно-транспортные происшествия.

Язык программирования: Python

Объем программы для ЭВМ: 54 МБ

РОССИЙСКАЯ ФЕДЕРАЦИЯ

**RU2025664579**

ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ
ГОСУДАРСТВЕННАЯ РЕГИСТРАЦИЯ ПРОГРАММЫ ДЛЯ ЭВМ

Номер регистрации (свидетельства):
2025664579
Дата регистрации: 04.06.2025
Номер и дата поступления заявки:
2025663699 04.06.2025
Дата публикации и номер бюллетеня:
04.06.2025 Бюл. № 6

Автор(ы):
Харрасов Камиль Раисович (RU)
Правообладатель(и):
Ордена Трудового Красного Знамени
федеральное государственное бюджетное
образовательное учреждение высшего
образования «Московский технический
университет связи и информатики» (RU)

Название программы для ЭВМ:

Автоматизированная информационная система оценки повреждений транспортного средства на основе видео контента

Реферат:

Современные системы обнаружения повреждений автомобилей сталкиваются с рядом ограничений, таких как снижение точности в условиях плохого освещения, трудности в распознавании мелких дефектов и высокая зависимость от качества изображений, что ограничивает их практическую применимость в реальных условиях эксплуатации. Предложенный метод обнаружения повреждений на базе нейро-сетевой модели использует оптимизированную архитектуру и специализированные датасеты, что позволяет значительно повысить точность и стабильность распознавания дефектов.

Язык программирования: Python

Объем программы для ЭВМ: 25 Мб

РОССИЙСКАЯ ФЕДЕРАЦИЯ

**RU2023681196**

ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ
ГОСУДАРСТВЕННАЯ РЕГИСТРАЦИЯ ПРОГРАММЫ ДЛЯ ЭВМ

Номер регистрации (свидетельства):
2023681196

Дата регистрации: 11.10.2023

Номер и дата поступления заявки:
2023669995 29.09.2023

Дата публикации и номер бюллетеня:
11.10.2023 Бюл. № 10

Контактные реквизиты:
нет

Автор(ы):

Харрасов Камиль Раисович (RU),

Мосева Марина Сергеевна (RU),

Городничев Михаил Геннадьевич (RU)

Правообладатель(и):

Ордена Трудового Красного Знамени
федеральное государственное бюджетное
образовательное учреждение высшего
образования «Московский технический
университет связи и информатики» (RU)

Название программы для ЭВМ:

Программа детектирования движущихся объектов при сложных погодных условиях

Реферат:

В настоящее время существующие системы детектирования имеют ряд недостатков, включая ограничения в точности, скорости реакции, способности действовать в сложных условиях, таких как сложные погодные условия или ограниченная видимость, что может привести к ошибкам распознавания. Программа позволяет осуществлять автоматизированное детектирование и классификацию объектов в условиях осложненной видимости, вызванной атмосферными осадками, в частности, различных транспортных средств и людей на стационарных и мобильных комплексах фото и видеофиксации.

Язык программирования: Python

Объем программы для ЭВМ: 12 МБ

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2026687102

**Программная реализация нейросетевого метода
обнаружения дорожных происшествий посредством
обработки видео контента**

Правообладатель: *Ордена Трудового Красного Знамени
федеральное государственное бюджетное
образовательное учреждение высшего образования
«Московский технический университет связи и
информатики» (RU)*

Автор(ы): *Харрасов Камиль Раисович (RU)*



Заявка № 2026686669

Дата поступления 01 сентября 2026 г.

Дата государственной регистрации

в Реестре программ для ЭВМ 07 сентября 2026 г.

*Руководитель Федеральной службы
по интеллектуальной собственности*

ДОКУМЕНТ ПОДПИСАН ЭЛЕКТРОННОЙ ПОДПИСЬЮ
Сертификат 00a570e477ad3d531b4b8818e75f29506
Владелец **Зубов Юрий Сергеевич**
Действителен с 04.09.2025 по 28.11.2026

Ю.С. Зубов

Приложение В. Акты о внедрении

«УТВЕРЖДАЮ»

Ректор ордена Трудового Красного
Знамени федерального государственного
бюджетного образовательного учреждения
высшего образования «Московский
технический университет связи и
информатики» (МТУСИ)



к.т.н., доцент Ерохин С.Д.

« 04 » 2026 г.

АКТ

об использовании результатов диссертационной работы Харрасова К.Р.
на тему: «Разработка нейросетевых методов обнаружения дорожно-транспортных
происшествий на основе обработки визуальной информации» в учебном процессе кафедры
«Математическая кибернетика и информационные технологии»

Комиссия в составе:

- заместителя проректора по учебной работе, Камаловой Гульназ Ильгамовны;
- заведующего кафедрой «Математическая кибернетика и информационные технологии», к.т.н., доц. Городничева Михаила Геннадьевича;

удостоверяет, что в учебном процессе кафедры «Математическая кибернетика и информационные технологии» (МКиИТ), при реализации образовательной программы «ТОП-ИИ: Инженерия систем искусственного интеллекта» (направление 09.03.01 «Информатика и вычислительная техника»), в рамках дисциплины «Системы машинного зрения», используются результаты диссертации Харрасова Камиля Раисовича, а именно: результаты анализа современных методов восстановления изображений в условиях затруднённой видимости; нейросетевой метод восстановления изображений от зашумлений, вызванных атмосферными осадками, обеспечивающий сохранение информативности визуальных признаков для последующей детекции объектов.

Указанные результаты включены в содержание дисциплины и используются при изучении методов цифровой обработки изображений, построения нейросетевых моделей машинного зрения и проектирования автоматизированных систем анализа видеоданных. Их внедрение способствует формированию у обучающихся практических компетенций в области разработки, оценки и применения интеллектуальных систем компьютерного зрения.

Заместитель проректора по учебной работе

Г.И. Камалова

Заведующий кафедрой МКиИТ

М.Г. Городничев



ОБЩЕСТВО С ОГРАНИЧЕННОЙ
ОТВЕТСТВЕННОСТЬЮ «ИНТЕЛСИТИ»

109316, РОССИЯ, Г. МОСКВА,
ВН. ТЕР. Г. МУНИЦИПАЛЬНЫЙ ОКРУГ ЮЖНОПОРТОВЫЙ,
ПР-КТ ВОЛГОГРАДСКИЙ, Д. 26, СТР. 1, ЭТАЖ 8, КОМ. 805

ИНН 7722019007 КПП 772201001
ОГРН 1227700178243

Исх. № 95/06

АКТ

о внедрении результатов диссертационной работы
Харрасова Камиля Раисовича, представленной на соискание ученой степени
кандидата технических наук

Настоящим актом подтверждается, что основные результаты диссертационного исследования Харрасова Камиля Раисовича «Разработка нейросетевых методов обнаружения дорожно-транспортных происшествий на основе обработки визуальной информации» в настоящее время используются в деятельности ООО «Интелсити», а именно:

- нейросетевой метод цифровой обработки визуальной информации с двухэтапным обнаружением дорожно-транспортных происшествий;
- формированием событий-кандидатов по динамике транспортных средств и их последующей нейросетевой классификацией;
- способ формирования признакового представления дорожной сцены на основе совместного анализа пространственных и спектральных характеристик с применением Fast Fourier Convolution, используемый при классификации событий-кандидатов;
- нейросетевой метод восстановления изображений от зашумлений, вызванных атмосферными осадками, со сниженной параметрической емкостью генератора, сохраняющий информативность визуальных признаков для последующей детекции объектов.

Применение указанных результатов обеспечивает полноту выявления дорожно-транспортных происшествий не ниже 0,88 на независимой тестовой выборке; при восстановлении изображений достигается структурное сходство не менее 0,78, а время обработки кадра сокращается вдвое и не превышает 29 мс.

Результаты используются при разработке и эксплуатации средств интеллектуального видеомониторинга дорожной обстановки и автоматизированного обнаружения дорожно-транспортных происшествий.

С уважением,

Генеральный директор



 / К.Ю. Мамедов



ООО «ТАГИР»

108840, г. Москва, г. Троицк, ул. Центральная, дом 5, ком. 21
 ОГРН 1247700638118, ИНН 7751331908, БИК 044030786,
 к/с 30101810600000000786, р/с 40702810832510001870
 Филиал «Санкт-Петербургский» АО «Альфа-Банк»

АКТ

о внедрении результатов диссертационной работы
 Харрасова Камиля Раисовича, представленной на соискание ученой степени
 кандидата технических наук

Настоящим актом подтверждается, что основные результаты диссертационного исследования Харрасова Камиля Раисовича «Разработка нейросетевых методов обнаружения дорожно-транспортных происшествий на основе обработки визуальной информации» в настоящее время используются в деятельности ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ "ТАГИР", а именно:

- архитектура автоматизированной информационной системы обнаружения дорожно-транспортных происшествий, обеспечивающая полный цикл получения видеопотоков, анализа дорожной сцены, формирования и хранения событий и передачи сообщений во внешние информационные системы;
- механизм адаптивного формирования последовательности обработки на основе сведений о погодных условиях с включением восстановления изображения только при ухудшении условий наблюдения;
- решения по асинхронному взаимодействию компонентов, хранению кадров в общей памяти без повторного копирования, пакетной обработке и горизонтальному масштабированию по числу видеопотоков.

Внедрение указанных результатов обеспечило масштабируемую обработку видеоданных: в исследованной конфигурации с разрешением 640×360 обрабатывались 50 видеопотоков при P95 = 85 мс и P99 = 91 мс; адаптивное включение восстановления исключает постоянное выполнение ресурсоемкой операции при нормальных условиях наблюдения.

Результаты используются при проектировании и эксплуатации автоматизированных информационных систем интеллектуального мониторинга дорожной обстановки и их интеграции с внешними информационными системами.

Руководитель



Губайдулина А.Н.