

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
«Уральский федеральный университет
имени первого Президента России Б.Н. Ельцина»

Институт радиоэлектроники и информационных технологий - РТФ
Базовая кафедра Аналитики больших данных и методов видеоанализа

ДОПУСТИТЬ К ЗАЩИТЕ ПЕРЕД ГЭК

/Зав. кафедрой Медведева М.А.

« 30 » мая 2025 г.

**ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
(МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ)**

**ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ ВИДЕОНАБЛЮДЕНИЯ ДЛЯ
ОБНАРУЖЕНИЯ АНОМАЛЬНОЙ АКТИВНОСТИ ЧЕЛОВЕКА**

Научный руководитель
к.т.н., доцент

подпись

А.Н. Медведев

Научный руководитель
к.ф.-м.н., доцент

подпись

М.А. Медведева

Студент группы
РИМ-230980

подпись

С.Д. Савченко

Екатеринбург
2025

РЕФЕРАТ

Тема магистерской диссертации:

Интеллектуальные системы видеонаблюдения для обнаружения
аномальной активности человека

Магистерская диссертация выполнена на 114 страницах, содержит 6 таблиц, 43 рисунка, 72 использованных источника.

Диссертация посвящена разработке алгоритма для автоматического обнаружения аномальной активности человека в видеопотоках, по предложенному многоуровневому подходу, который сочетает методы глубокого обучения, частотного анализа и статистического моделирования. Актуальность исследования обусловлена ростом угроз общественной безопасности и ограничениями традиционных систем видеонаблюдения, зависящих от человеческого фактора и неспособных эффективно обрабатывать сложные сцены с большим количеством людей, шумом и перекрытиями объектов. Современные камеры генерируют огромные объемы данных, но их ручной анализ становится невозможным из-за ограниченных возможностей операторов, что требует внедрения автоматизированных решений, способных работать в реальном времени с минимальными трудозатратами.

Цель работы заключается в разработке алгоритма обработки видеопотока с камер видеонаблюдения для оптимизации процесса обнаружения аномальной активности человека. Для достижения цели были сформулированы и решены задачи: обосновать выбор методического подхода к обнаружению аномальной активности человека на видеопотоках, разработать модель обнаружения аномальной активности человека на видеопотоках с камер видеонаблюдения и разработать алгоритм обработки видеоданных для выявления аномальной активности человека с использованием предложенных моделей.

Объектом исследования выступают подходы и методы обнаружения аномальной активности человека в системах видеонаблюдения. Предметом исследования являются Модели и алгоритмы обработки видеоданных для выявления аномальной активности человека. Основное внимание уделено интеграции различных методов в один конвейер обработки видеопотока для улучшения эффективности обнаружения аномальных событий.

Научная новизна работы выражается в предложении многоуровневого подхода к обнаружению аномальной активности человека в видеопотоке, который включает методы глубокого обучения, частотного анализа и статистического моделирования, отличающийся интеграцией графов событий и LOF-анализа, что в итоге позволяет повысить устойчивость к фоновым помехам и перекрытиям объектов.

Практическая значимость заключается в создании алгоритма, который способен автоматически анализировать видеопотоки, обнаруживать и классифицировать аномальное поведение человека, в автономном режиме

Экономическая эффективность проявляется в сокращении расходов на персонал за счет автоматизации анализа видеопотока, повышении скорости реакции на инциденты и снижении ущерба от угроз.

СОДЕРЖАНИЕ

РЕФЕРАТ	5
ВВЕДЕНИЕ	8
1 АНАЛИЗ СУЩЕСТВУЮЩИХ МЕТОДОВ ОБНАРУЖЕНИЯ АНОМАЛЬНОЙ АКТИВНОСТИ ЧЕЛОВЕКА В ВИДЕОАНАЛИТИЧЕСКИХ СИСТЕМАХ	13
1.1 Актуальность исследования в сфере видеонаблюдения	13
1.2 Существующие подходы и методы обнаружения аномальной активности	16
1.3 Применение комбинированного подхода для обнаружения аномальной активности человека	43
1.4 Постановка задачи управления системой	45
1.5 Результаты и выводы первой главы	50
2 РАЗРАБОТКА МОДЕЛИ ОБНАРУЖЕНИЯ АНОМАЛЬНОЙ АКТИВНОСТИ ЧЕЛОВЕКА С ИСПОЛЬЗОВАНИЕМ МНОГОУРОВНЕВОГО ПОДХОДА	51
2.1 Описание разрабатываемой модели для обнаружения аномальной активности человека	51
2.2 Анализ существующих моделей для предобработки изображений и извлечения признаков	52
2.3 Выбор исходного набора данных для обучения модели	57
2.4 Результаты и выводы второй главы	66
3 ПРОГРАММНАЯ РЕАЛИЗАЦИЯ АЛГОРИТМА ОБРАБОТКИ ВИДЕОДАНЫХ ДЛЯ ВЫЯВЛЕНИЯ АНОМАЛЬНОЙ АКТИВНОСТИ ЧЕЛОВЕКА	69
3.1 Предпроектное обследование предприятия для внедрения модели в системы видеонаблюдения	69
3.2 Определение требований для внедрения системы	72
3.3 Цифровой двойник процесса мониторинга с использованием алгоритмов автоматического обнаружения аномальных событий	74
3.4 Описание методологии разработки	77
3.5 Программная реализация алгоритма для обнаружения аномальной активности	79
3.6 Тестирование разработанной модели	98
3.7 Результаты и выводы третьей главы	101
ЗАКЛЮЧЕНИЕ	103
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	105
ПРИЛОЖЕНИЕ А	114

ВВЕДЕНИЕ

Актуальность темы исследования. В условиях роста угроз общественной безопасности, таких как кражи, вандализм и террористические акты, вопросы обеспечения безопасности становятся особенно значимыми. Для обеспечения безопасности в современном мире широко применяются системы видеонаблюдения (CCTV). Такие системы используются в городской инфраструктуре, транспорте и частных учреждениях, однако эффективность определения аномальных событий ограничена зависимостью от человеческого фактора.

Решение задачи повышения эффективности мониторинга и снижения влияния человеческого фактора требует внедрения автоматизированных методов анализа видеопотоков, способных обнаруживать аномалии в реальном времени без участия операторов. Существующие решения, основанные на анализе видеопотоков, демонстрируют эффективность лишь в упрощенных сценариях, сталкиваясь с рядом проблем в динамичной среде. Например, статистические модели не учитывают взаимодействие между объектами, а простые детекторы движения часто реагируют на незначительные изменения, вызывая ложные срабатывания.

Интеллектуальные системы видеонаблюдения, основанные на современных алгоритмах машинного обучения, могут стать ключевым решением для повышения уровня безопасности. Такие системы позволяют автоматически выявлять подозрительное поведение, классифицировать типы событий и формировать отчеты, минимизируя участие человека в процессе. Таким образом, исследование и разработка интеллектуальных систем видеонаблюдения актуальны и имеют высокую практическую ценность в обеспечении общественной безопасности.

Степень разработанности темы исследования. Проблемы автоматического обнаружения аномальной активности в видеопотоках изучаются уже более двух десятилетий. Первые подходы были основаны на использовании статистических моделей и пространственно-временного анализа,

которые позволяли выявлять отклонения от нормального поведения. Однако такие методы оказались недостаточно устойчивыми к шуму и сложным сценам.

С развитием технологий машинного обучения начали применяться более продвинутое подходы: сверточные нейронные сети (CNN), автоэнкодеры, SVM и ансамблевые методы. Одним из первых шагов стало использование VGGNet и ResNet для извлечения признаков из видеок кадров, что позволило повысить точность анализа. Также были предложены гибридные модели, объединяющие преимущества различных методов — например, комбинация CNN с графами событий или частотный анализ с GMM.

Особое внимание уделялось разработке систем, не зависящих от трекинга объектов, чтобы повысить устойчивость к перекрытиям и шуму. Подходы, основанные на анализе движений и динамике толпы, также получили развитие, особенно в работах по моделированию социальных взаимодействий и когнитивных процессов.

Несмотря на достигнутые успехи, существующие методы все еще сталкиваются с рядом ограничений: высокая чувствительность к внешним условиям, необходимость большого объема размеченных данных, ограниченная адаптивность к новым типам аномалий и сложность масштабирования. Особенно остро стоит вопрос создания универсальной системы, способной функционировать в реальном времени с минимальным количеством ложных срабатываний и высокой точностью распознавания.

Объект исследования – Подходы и методы обнаружения аномальной активности человека в системах видеонаблюдения.

Предмет исследования – Модели и алгоритмы обработки видеоданных для выявления аномальной активности человека.

Целью данного исследования является разработка алгоритма обработки видеопотока с камер видеонаблюдения для оптимизации процесса обнаружения аномальной активности человека. Для достижения этой цели исследования были поставлены следующие задачи:

1. Обосновать выбор методического подхода к обнаружению аномальной активности человека на видеопотоках.
2. Разработать модель обнаружения аномальной активности человека на видеопотоках с камер видеонаблюдения.
3. Разработать алгоритм обработки видеоданных для выявления аномальной активности человека с использованием предложенных моделей.

Научная новизна исследования:

1. Предложен многоуровневый подход к обнаружению аномальной активности человека в видеопотоке, включающий методы глубокого обучения, частотного анализа и статистического моделирования, отличающийся интеграцией графов событий и LOF-анализа, позволяющей повысить устойчивость к фоновым помехам и перекрытиям объектов.

2. Разработана модель обнаружения аномальной активности человека на видео, по предложенному многоуровневому подходу, включающая предобработку с векторным и частотным преобразованием данных, отличающуюся возможностью мультимодального анализа данных, позволяющей повысить точность анализа динамики сцены за счет частотного представления.

3. Разработан алгоритм обработки видеопотока, включающий многоуровневую структуру обработки входного видеопотока для обнаружения аномальной активности человека, отличающийся интеграцией методов в единый конвейер обработки видеопотока, позволяющий повысить точность обнаружения аномального события.

Методология и методы исследований. В работе используется методология системных исследований, которая предполагает комплексный подход к изучению сложных технических систем, в данном случае — системы обнаружения аномальной активности человека в видеопотоке. Такой подход позволил рассмотреть задачу с разных ракурсов от уровня отдельного кадра до

масштабирования на всю инфраструктуру видеонаблюдения. Для достижения поставленных целей в исследовании использовались следующие методы:

- Анализ научной литературы и существующих алгоритмов.
- Методология разработки Agile.
- Экспериментальная проверка эффективности различных моделей (VGGNet, FFT, GMM, LOF, ансамблевые методы) на датасете UCF-Crime, включающем 14 типов аномальных событий и «нормальные» видеозаписи.

Положения, выносимые на защиту:

1. Предложен подход к обнаружению аномальной активности человека в видеопотоке, **включающий** методы глубокого обучения, частотного анализа и статистического моделирования, **отличающийся** интеграцией графов событий и LOF-анализа, **позволяющей** повысить устойчивость к фоновым помехам и перекрытиям объектов.

2. Разработана модель обнаружения аномальной активности человека на видео, по предложенному многоуровневому подходу, **включающая** предобработку с векторным и частотным преобразованием данных, **отличающуюся** возможностью мультимодального анализа данных, **позволяющей** повысить точность анализа динамики сцены за счет частотного представления.

3. Разработан алгоритм обработки видеопотока, **включающий** многоуровневую структуру обработки входного видеопотока для обнаружения аномальной активности человека, **отличающийся** интеграцией методов в единый конвейер обработки видеопотока, **позволяющий** повысить точность обнаружения аномального события.

Достоверность научных результатов обеспечивается корректным применением современных методов машинного обучения, использованием проверенных наборов данных (UCF-Crime, Violent-Flows, ShanghaiTech Campus), а также проведением сравнения результатов обнаружения аномальных событий с известными решениями, которые показали улучшение по метрикам точности.

Личный вклад заключается в постановке задачи управления, проведении анализа предметной области, формулировании целей и задач, а также разработке модели интеллектуальной системы для обнаружения аномального поведения человека с помощью видеокамер наружного наблюдения. **Апробация результатов.** Основные результаты представлены на международных научных конференциях: XIX Международная конференция «Российские регионы в фокусе перемен», III Международная конференция «Цифровизация и цифровая трансформация экономики и общества: современные тенденции и инструменты»

Диссертационная работа состоит из введения, 3 глав, заключения. Объем 111 страниц, содержит 5 таблиц, 43 рисунка, количество использованных литературных источников составляет 72.

1 АНАЛИЗ СУЩЕСТВУЮЩИХ МЕТОДОВ ОБНАРУЖЕНИЯ АНОМАЛЬНОЙ АКТИВНОСТИ ЧЕЛОВЕКА В ВИДЕОАНАЛИТИЧЕСКИХ СИСТЕМАХ

1.1 Актуальность исследования в сфере видеонаблюдения

В условиях современного мира, которые характеризуются ускоренным технологическим прогрессом и усложнением социальных связей, угрозы общественной безопасности приобретают глобальный характер. Традиционные методы защиты, такие как физическая охрана или простейшие системы видеонаблюдения, использовались в условиях, когда угрозы были локальными и легко идентифицировались. Сегодня все больше стран сталкивается с новыми тенденциями в виде урбанизации общества [1], которая приводит к перенасыщению общественных пространств в городах, и всеобщей цифровизации, которая дает возможность быстрого распространения информации из-за чего преступные схемы становятся более координированными, а высокая плотность населения и скорость распространения информации приводят к тому, что локальные конфликты могут быстро эскалировать в массовые. Перечисленные тенденции в совокупности со сложной геополитической обстановкой в мире делают сферу обеспечения безопасности наиболее значимой. Проблема защиты как публичных, так и частных пространств от различных угроз, включая кражи, вандализм и террористические акты, требует внедрения технологий мониторинга и наблюдения [2]. Одним из ключевых решений обеспечения безопасности стали системы видеонаблюдения (CCTV), которые сегодня применяются в различных местах — от городских улиц до учреждений и частных домов [3].

По данным Precedence Research [4] мировой рынок систем видеонаблюдения оценивается в 51 млрд долларов США, к 2030 году объем может вырасти до 130 млрд долларов США.

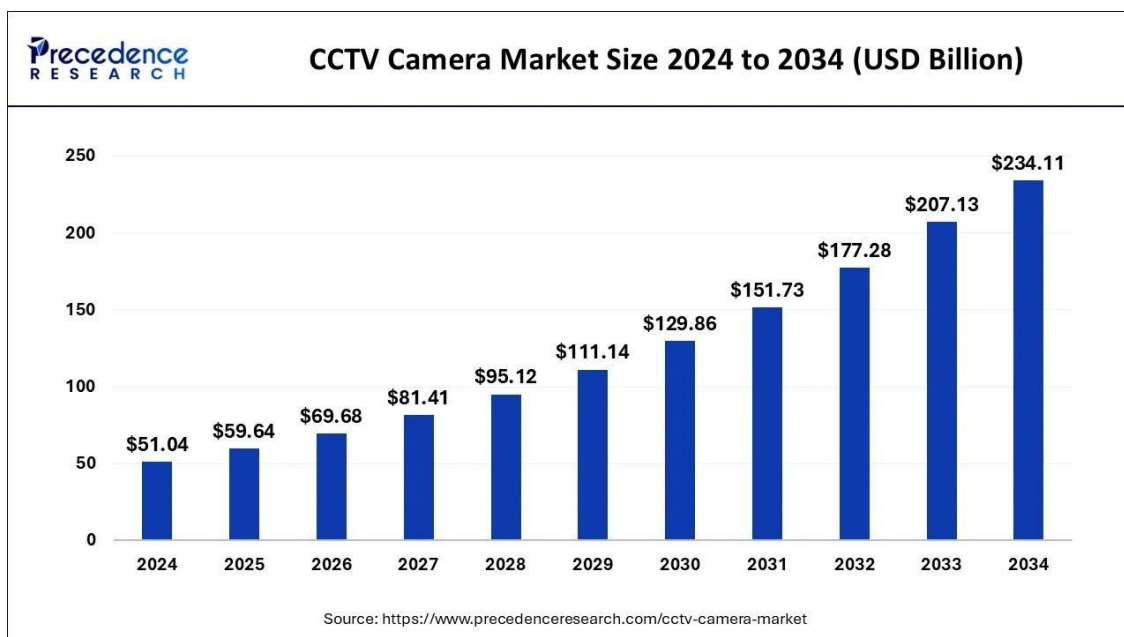


Рисунок 1 – Прогноз объема рынка систем видеонаблюдения

Высокая потребность в системах видеомониторинга обуславливается рядом факторов, основным из которых является усиление мер по обеспечению общественной безопасности - профилактика и раскрытие преступлений, терактов и прочих негативных процессов. Помимо этого, повышению интереса к CCTV также способствуют экономические факторы, так внедрение видеонаблюдение снижает риски потерь от преступлений и позволяет снизить затраты на физическую охрану. Значимую роль в развитии рынка играют государства, которые развивают концепции умных городов, например, система умного города, запущенная в Сингапуре, использует более 100 тысяч камер видеонаблюдения и к 2030 году их количество достигнет 200 тысяч [5]. В России также развивается сфера видеонаблюдения, так в 2024 году госкомпания Российской Федерации увеличили объем закупок систем видеомониторинга на 131% по сравнению с предыдущим годом.

В данный момент наибольшую популярность получили камеры видеонаблюдения с автоматической детекцией движения, такое решение позволяет более эффективно проводить процесс мониторинга, однако детекторы движения реагируют на любое изменение в кадре, из-за чего не подходит для мониторинга сцен, где движение людей считается нормальным поведением.

Наиболее развитый контроль и автоматизация принятия решений реализован в системах наблюдения за дорожным движением, где система определяет нарушения без участия человека [6]. Несмотря на достижения в автоматизации процессов видеонаблюдения в дорожном движении, контроль за поведением человека в общественных местах все еще осуществляется с участием операторов. В таких системах операторам приходится вручную отслеживать множество камер, что не только требует значительных трудозатрат, но и создает потенциальные риски ошибок и упущений. Человеческий фактор в этом процессе может привести к неправильной интерпретации ситуации, занижению или завышению уровня угрозы, а также к пропуску важных событий, которые требуют немедленного реагирования [7]. В условиях высокой нагрузки на операторов и большого объема информации, поступающего одновременно с нескольких камер, эффективность мониторинга серьезно снижается. Утомление, рассеянность или недостаточный уровень квалификации специалистов могут негативно сказаться на качестве контроля [8]. Это создает необходимость в квалифицированном персонале, что в свою очередь увеличивает затраты и усложняет организацию охраны.

Таким образом, можно выделить несколько групп основных проблем использования систем видеонаблюдения в текущей ситуации:

- Проблемы, возникающие из-за современных тенденций.
- Проблема человеческого фактора.
- Технические проблемы.

К первой группе проблем можно отнести следующие примеры:

- Увеличение числа видеокамер при ограниченных способностях операторов мониторинга.
- Высокая динамика отслеживаемых сцен при низкой адаптивности используемых систем.

К проблемам, возникающим из-за человеческого фактора:

- Субъективность интерпретации, то есть операторы могут ошибочно классифицировать события из-за усталости, стресса или недостатка опыта.

- Снижение эффективности в зависимости от времени. Внимание человека рассеивается со временем, то есть эффективность распознавания аномальных событий в начале рабочей смены может быть ниже, чем в конце.

- Зависимость от квалификации персонала, которая влечет за собой сложности с масштабируемостью.

Существующие системы обладают недочетами, из-за которых возникают следующие проблемы:

- Отсутствие распознавания паттернов поведения. Камеры оснащенные детекторами движения реагируют на любые движения в кадре, даже если поведение является нормой.

- Неспособность классификации события и создания автоматического отчета.

Описанные тенденции подтверждают важность сферы видеонаблюдения, разработки и применения автоматизированных систем, которые способны обрабатывать и анализировать видеоданные в реальном времени, выявляя поведение, которое может указывать на угрозы или нарушения, без постоянного человеческого вмешательства. Это позволит существенно разгрузить операторов и повысить общую эффективность систем видеонаблюдения [9]. Более того, интеграция таких решений не только ускоряет процесс реагирования на инциденты, но и обеспечивает более высокую степень точности в распознавании критически важных событий.

1.2 Существующие подходы и методы обнаружения аномальной активности

Детекция аномального поведения с помощью камер видеонаблюдения стала темой для исследований, охватывающих широкий спектр подходов и методов. Для обнаружения аномальной активности используются различные научные подходы:

- Подход, основанный на создании социологических моделей.

- Моделирование когнитивных процессов, которые формализуют эвристики и поведенческие модели.

- Гибридный подход Мохаммади который внедряет набор простых поведенческих эвристик в физические уравнения.

- Подходы, не требующие трекинга объектов.

- Подходы, основанные на пространственно-временных данных.

- Подходы, использующие алгоритмы машинного обучения.

К первой категории научного подхода относятся работы, напрямую использующие физические аналогии для моделирования поведения толпы. Наиболее ярким примером является модель социальной силы (SFM) Хелбинга и Молнара [10], модель, вдохновленная ньютоновской механикой, описывает взаимодействие толпы, была использована для обнаружения аномального поведения людей с помощью компьютерного зрения в работе Мехрана Р. [11]. Свои варианты социологических моделей для обнаружения аномалий предлагали Вильям Зенг в работе по обнаружению нарушений правил пешеходами на перекрестках [12], Даниэль Париси в работе по моделированию пешеходных потоков [13], а также Франческо Занлунго для прогнозирования столкновений [14]. Использование этого научного подхода не оптимально, так как центральный принцип всех моделей – описание различных сценариев скопления людей путем калибровки набора физических сил на основе эмпирических наблюдений. Исследования в области социальной психологии [15-16] утверждают, что модели слишком упрощены для описания сложного поведения толпы. Также поведение толпы имеет слабую способность к обобщению, а это означает, что модель, рассчитанная на основе набора эмпирических наблюдений, часто может не соответствовать другому набору наблюдений.

Для преодоления ограничений социологических моделей был разработан подход, использующий моделирование когнитивных процессов и выделение поведенческих эвристик. В основе научного подхода лежат работы Лауры Мартиггон [17], Джона Хатчинсона [18] и Ральфа Хертвига [19]. Эвристические

правила, например, «следуй за лидером», «избегай столкновений» внедряются в физические уравнения для улучшения прогнозирования реакций в стрессовых ситуациях. Примером служит работа Fight or Flight [20], которая интерпретирует резкие движения как проявления агрессии или паники. Такие методы полезны для анализа ранних признаков аномалий, но сталкиваются с проблемами из-за низкой применимости в многолюдных зонах, так как эвристики направлены на анализ индивидуальных реакций.

Новый научный подход представил Садег Мохаммади, предложив гибридную структуру для обнаружения аномального поведения [21]. Гибридная структура подразумевает объединение моделей SFM с поведенческими эвристиками, такой подход особенно эффективен для анализа давки и панических реакций, так как снижает количество ложных срабатываний за счет учета социальных и физических факторов. Однако его реализация требует настройки параметров под конкретные сценарии и чувствительна к ошибкам трекинга.

Четвертый подход используется для ситуаций, когда сложно получить надежные данные о перемещении. Цуй Ксини предложил использовать методы на основе гистограмм [22] Вей-Ксин Ли применил смеси динамических текстурных моделей [23]. Тимоти Хоспедальс использует тематическое моделирование а Янгванг Жу скрытые марковские модели (НММ) [24]. Предложенные методы не используют информацию о перемещении объектов, анализируя вместо этого данные более высокого уровня, такой подход не имеет возможности анализа данных о взаимодействии между людьми, что снижает информативность.

Более современным и совершенным подходом считается – анализ пространственно-временных признаков [25-28].

В работе Фабио де Сойза [29] использовался пространственно-временной детектор точки интереса (STIP) и дескриптор наряду с машинами линейных опорных векторов (SVM). Энрико Ньевас использовал гистограмму градиентов

(HOG) и Motion SIFT (MoSIFT) для обнаружения аномальных событий [30]. Другие подходы выводили закономерности локального движения из оптических потоков, например Беркан Солмаз проанализировал потоки движения для идентификации конкретного набора простых моделей поведения толпы. К недостаткам данного подхода можно отнести высокую вычислительную нагрузку

Самым современным подходом считается использование машинного обучения, который также содержит в себе методы, основанные на пространственно-временных данных, которые имеют важное значение для анализа динамичных сцен [31]. В методах данного научного подхода используются модели машинного и глубоко обучения, таких как сверточные нейронные сети (CNN), поддерживающие векторные машины (SVM) и сети долговременной краткосрочной памяти (LSTM) [32]. Эти методы доказали свою эффективность в контексте обнаружения аномальных событий в видеопотоках. Сравнение подходов к обнаружению аномальной активности человека на видео показано в таблице 1

Таблица 1 - Сравнение подходов для обнаружения аномальной активности

Подход	Преимущества	Недостатки
Моделирование социологических моделей	Учитывает взаимодействие между людьми. Прогнозирует развитие сценариев давки и панической реакции.	Требует детальной настройки под конкретные сценарии. Чувствителен к ошибкам трекинга. Сложно адаптируется к новым условиям. Низкая обобщающая способность.
Моделирование когнитивных процессов и эвристик	Объясняет индивидуальные реакции в стрессовых ситуациях. Позволяет прогнозировать ранние признаки агрессии или паники.	Неэффективен в многолюдных сценах. Сфокусирован на индивидуальном поведении

Продолжение таблицы 1 - Сравнение подходов для обнаружения аномальной активности

Гибридный подход Мохаммади	Повышает устойчивость к ложным срабатываниям. Учитывает как физические, так и социальные факторы. Эффективен при анализе давки и панических реакций	Требует ручной настройки параметров. Зависит от качества трекинга объектов. Ограниченная универсальность.
Подходы, не требующие трекинга объектов	Работает без точного трекинга. Устойчив к перекрытиям объектов. Может применяться в условиях высокой плотности людей.	Низкая информативность из-за отсутствия данных о взаимодействиях. Ограниченная точность распознавания. Мало пригоден для классификации сложных событий.
Пространственно-временные методы	Высокая чувствительность к локальным изменениям. Учитывает временные зависимости. Применим к разнообразным сценариям.	Высокая вычислительная нагрузка. Чувствителен к шуму и фоновым помехам. Ограниченная масштабируемость.
Алгоритмы машинного и глубокого обучения	Автоматическое извлечение признаков. Устойчив к шуму и перекрытиям объектов. Работа с немаркированными данными.	Требует вычислительных ресурсов. Зависит от качества данных. Требует начального этапа обучения. Сложность интерпретации нейросетевых решений.

В результате сравнения подходов к обнаружению аномальной активности, наиболее потенциальным является подход с использованием машинного обучения.

Одним из значимых направлений является использование пространственно-временных подходов, позволяющих интегрировать информацию о движении и контексте. Например, Илан Ван [33] предлагает метод, который позволяет обучать модели на немаркированных данных. Этот подход достигает двух

ключевых целей: прогнозирования значимости и обнаружения аномалий. Для этого используется быстрое преобразование Фурье (FFT) для преобразования сигналов в частотную область, что затем анализируется с применением теоремы Парсеваля. Общая схема работы метода Илана Вана представлена на рисунке 2



Рисунок 2 – Схема алгоритма Илана Вана

На этапе предобработки видеопоток разбивается на последовательность кадров и выделяются векторы движения объектов методом оптического потока.

Затем на втором этапе преобразования в частотную область, временные ряды векторов движения обрабатываются с помощью быстрого преобразования Фурье (FFT), что позволяет выделить частотные компоненты, характерные для "нормального" поведения. FFT - алгоритм, который преобразует сигнал из временной области в частотную, разбивая его на гармонические компоненты. Высокочастотные гармоники (например, 10–30 Гц) могут указывать на резкие движения (падения, насилие) а низкочастотные (0–5 Гц) соответствуют плавным изменениям (например, движение толпы). Данный алгоритм применяется для моделирования нормальных паттернов на основе исторических данных, то есть обнаружение аномалий возможно на неразмеченных данных.

На третьем этапе происходит анализ энергии сигнала с использованием теоремы Парсеваля. Система рассчитывает энергию сигнала во временной области (движение объектов) и частотной (спектр от FFT). При значительном расхождении между ними, например, энергия во временной области резко возрастает, а в частотной — нет, событие классифицируется как аномалия.

На четвертом этапе система определяет зоны, требующие внимания основе анализа частотных паттернов.

К преимуществам метода Илана Вана можно отнести:

- Возможность обнаружения аномалий без разметки данных за счет обучения на нормальных паттернах.
- Устойчивость к шуму за счет фильтрации фоновых изменений.

Основным недостатком такого подхода является требуемая вычислительная мощность. Для обработки кадров с частотой более 24 кадров в секунду,

Динеш Сингх [34] использует метод построения графов для анализа видео, где пространственно-временные точки интереса представлены как вершины графа, а связи между объектами – как ребра. Этот подход позволяет применять нечеткие функции принадлежности для нахождения аномалий и классификации поведения с помощью графового ядра SVM. Общее описание алгоритма Динеша Сингха с построением графов представлено на рисунке 3



Рисунок 3 – метод Динеша Сингха

На первом этапе извлечения ключевых точек вершины графа выделяются с помощью метода оптического потока. Каждая вершина соответствует локализованной активности, например, резкому движению руки или изменению скорости.

Далее строятся связи между ключевыми точками если между ними существует временная или пространственная корреляция, например два человека, двигающихся в одном направлении, формируют ребро с весом, пропорциональным их синхронности.

Анализ графов происходит посредством графового ядра SVM, которое вычисляет степень принадлежности к «норме» или «аномалии», например, от 0 до 1 для каждой вершины и ребра.

К преимуществам данного метода относятся:

- Моделирование взаимодействий между объектами – в данном подходе учитываются глобальные связи между разными объектами, которые недоступны для методов, анализирующих отдельные движения.
- Устойчивость к неопределенностям в данных – благодаря использованию нечеткой логики метод позволяет работать с неопределенными данными, например при перекрытии объектов.
- Гибкость графового подхода – метод можно адаптировать к разным сценариям через изменение правил построения ребер.

Недостатками подхода является высокая вычислительная сложность, а также требовательность к качеству видеопотока.

Метод, предложенный Фейпингом Ли [35], основан на вероятностном моделировании нормального поведения через гауссовские смеси (GMM), что позволяет выявлять аномалии без необходимости отслеживания отдельных объектов. В отличие от одномерных гауссовых моделей, GMM учитывает многомодальные распределения, характерные для динамичных сцен, например,

разные направления движения в толпе. Алгоритм состоит из 4 условных этапов, представленных на рисунке 4

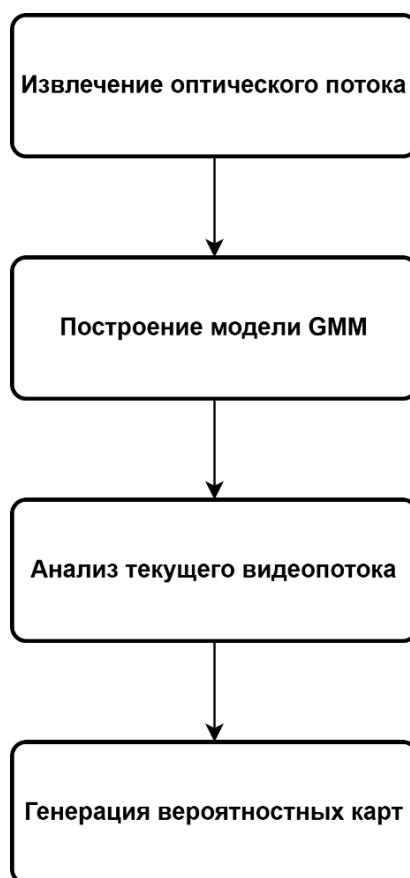


Рисунок 4 – Метод Фейпинга Ли

На первом этапе происходит извлечение оптического потока, с помощью алгоритма Лукаса-Канаде, который выделяет ключевые точки на кадре (углы и контуры) и рассчитывает векторы движения между кадрами. Данный алгоритм обладает линейной сложностью, то есть имеет низкие требования к системе.

На этапе построения модели GMM формируется вероятностная модель, представляющая нормальное распределение векторов оптического потока. GMM представляет собой набор гауссовых распределений, каждое из которых моделирует определенный тип движения (например, движение влево, вправо, статичное состояние)

При анализе текущего видеопотока на третьем этапе, для каждого кадра рассчитывается вероятность $p(x)$ того, что движение соответствует норме. Если

рассчитанная вероятность меньше порога чувствительности, событие определяется как аномальное.

К преимуществам данного метода относятся:

- Моделирование многомодальных распределений, то есть GMM учитывает разные типы движений, например, разные направления в толпе.
- Устойчивость к фоновым помехам за счет вероятностного подхода система игнорирует незначительные изменения.
- Метод не требует разметки данных, так как обучается на нормальных видеопотоках.
- Линейная сложность алгоритма позволяет работать в реальном времени при меньших требуемых аппаратных ресурсах.

Из недостатков метода можно выделить:

- Чувствительность к шуму и освещению, так как изменение условий искажают векторы оптического потока снижая точность.
- Ограниченная адаптивность к сложным сценам - В толпе с перекрытиями объектов GMM может пропустить аномалии из-за усреднения векторов оптического потока.
- Необходимость ручной настройки порога вероятности нормального события.

Метод, предложенный Чжаохуэем Ло [36], основан на пространственно-временном моделировании с использованием функции Smoothing Motion Patterns (SMP) и гауссовских смесей (GMM), который завершается классификацией с помощью SVM. Основная цель — выявлять аномалии в видеопотоках за счет сглаживания движений и последующего анализа отклонений от "нормальных" паттернов. Упрощенное графическое представление алгоритма представлено на рисунке 5



Рисунок 5 – Метод Чжаохуэя Ло

На этапе извлечения признаков применяется SMP (Smoothing Motion Patterns) с целью сглаживания шумовых компонент в видеопотоке и выделения устойчивых паттернов движения. SMP анализирует последовательность кадров для выделения гладких траекторий объектов и использует фильтрацию временных рядов для подавления случайных колебаний (вибрации, тени). Также для каждого объекта выделяются пространственно-временные признаки:

- Скорость, обозначающая изменение координаты x за время t .
- Направление движения.
- Ускорение.

Из этих параметров формируются признаки $x=(v,\theta,a)$ для дальнейшего анализа.

При моделировании нормальных паттернов используется ранее описанный GMM, который представляет данные как смесь гауссовых распределений. В результате GMM формирует вероятностные карты "нормального" поведения, где зоны с низкой вероятностью интерпретируются как потенциальные аномалии.

На заключительном этапе происходит обнаружение аномалий с помощью SVM, то есть события классифицируются как аномальные или нормальные на основе признаков, извлеченных с помощью SMP и GMM. На вход SVM подаются векторы $x(\text{smp})$, дополненные вероятностями $p(x)$ от GMM, после чего SVM строит разделяющую гиперплоскость, которая максимизирует расстояние между классами.

К преимуществам метода относятся:

- Сниженное влияние выбросов за счет того, что SMP фильтрует кратковременные помехи (например, толчки в толпе), а GMM усредняет распределение движений.
- Высокая точность в стабильных сценах.

Но метод имеет недостатки, например:

- Сниженное влияние выбросов за счет того, что SMP фильтрует кратковременные помехи (например, толчки в толпе), а GMM усредняет распределение движений.
- Высокая точность в стабильных сценах.
- Чувствительность к качеству видеопотока, так как SMP может не справляться с фильтрацией, что приводит к ложным срабатываниям.
- Резкое снижение точности моделирования поведения при перекрытии объектов.
- Зависимость от разметки данных для обучения SVM.
- Высокая ресурсоемкость SMP в связке с GMM требует около 1 секунды на обработку одного кадра, что делает метод непригодным для анализа в реальном времени.

Далее, Юань Юань [37] рассматривает стратегию обнаружения аномалий, основанную на статистическом детекторе. Метод анализирует видеоданные, где каждый кадр или последовательность кадров рассматривается как набор паттернов, представленных в виде гистограмм. Основная идея заключается в

том, чтобы выявлять аномалии как статистически значимые отклонения от нормального распределения признаков, при этом минимизируя влияние шума.

Алгоритм Юаня Юаня в общем виде представлен на рисунке 6

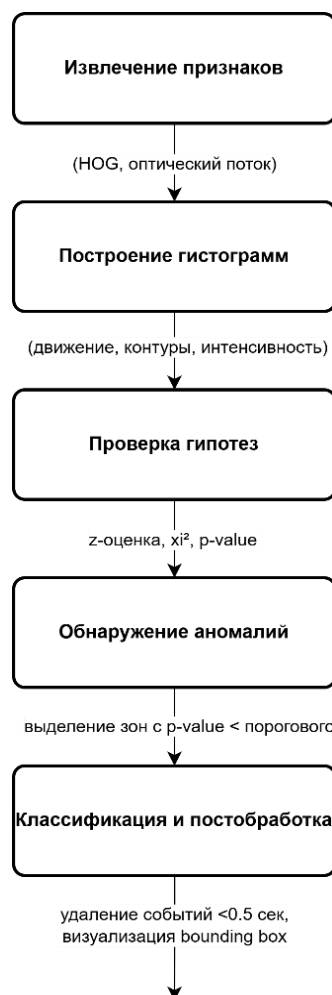


Рисунок 6 – Метод Юаня Юаня

На этапе извлечения признаков входной видеопоток преобразуется в набор гистограмм, описывающих движение, контуры и изменения в сцене. Эти данные становятся основой для построения статистических моделей "нормы". Признаки извлекаются в гистограммы трех типов:

- **Гистограммы градиентов (HOG)** анализируют направленность градиентов в локальной области кадра, затем гистограмма разбивается на бины каждый бин соответствует диапазону направлений, например, 0–180°.

- **Гистограммы движения** – строятся по векторам оптического потока по направлениям и скоростям движения объектов.

- **Гистограммы интенсивности изменений**, которые анализируют разницу между соседними кадрами.

На следующем этапе создаются эталонные распределения признаков, характерные для нормального поведения, то есть система анализирует видеопотоки с нормальными сценами, например, движение толпы без аномалий. После чего для каждого типа признака (HOG, движение, интенсивность) строится гистограмма нормы.

Для определения отклонения параметров от нормы используются статистические критерии. Для каждого кадра формируется статистическая оценка аномальности через z-оценки, χ^2 и p-value, где:

- Z-оценка - указывает на аномалию в соответствующей области кадра.
- χ^2 - вычисляет соответствие текущей гистограммы норме.
- P-value - вероятность наблюдения текущих данных при условии нормы.

Аномалии обнаруживаются на основе статистических тестов. Бины с высоким значением Z-оценки выделяются как подозрительные локальные отклонения. На глобальном уровне анализируются χ^2 значения, при превышении которых система классифицирует кадр как аномальный. Для каждого кадра создается карта, где яркость пикселей отражает степень отклонения от нормы. В результате кадры с аномалиями фильтруются, а их координаты передаются на этап классификации.

На этапе классификации статистические отклонения переводятся в интерпретируемые человеком термины, такие как насилие, падение и другие.

В итоге метод показывает следующие преимущества:

- Сниженное влияние шума благодаря усреднению в бинах.
- Автоматическая адаптация порогов чувствительности в зависимости от условий.

- Обнаружение аномалии без разметки данных, используя статистические тесты (z -оценка, χ^2).

Однако метод сталкивается с ограничениями в виде неспособности анализировать сложные сцены из-за упрощенных статистических моделей.

Метод, предложенный Дэн Сюем [38], основан на разделении видеопотока на движение и статику с помощью структуры двойного слияния, за которой следует автоматическая классификация аномалий через автоэнкодеры и SVM. Основная цель — повысить точность обнаружения аномалий за счет выделения динамических компонентов, игнорируя фоновые изменения. Схема метода представлена на рисунке 7



Рисунок 7 – Метод Дэна Сюема

Первым этапом в методе Дэна Сюема является разделение признаков с помощью Two-stream fusion, которое разделяет видеопоток на движущиеся и статичные элементы, чтобы сосредоточить анализ на динамике. Первый поток –

динамические движения получается через оптический поток. Анализ статических данных происходит с помощью сверточных нейронных сетей.

Два потока комбинируются в вектор x_{fusion} , который служит входом для автоэнкодера.

Автоэнкодер снижает размерность данных и снижает влияние шума, сохраняя при этом информативные признаки. Энкодер преобразует вектор x_{fusion} в сжатое представление $z \in \mathbb{R}^{128}$. Декодер выполняет обратную роль реконструкции данных минимизируя потери:

$$L_{reconstruction} = |x_{fusion} - \hat{x}_{fusion}|$$

Где \hat{x}_{fusion} - восстановленный вектор.

В результате на выходе автоэнкодера получается 128-мерный вектор z , устойчивый к шуму и фоновым изменениям.

На этапе классификации с использованием SVM аномалии идентифицируются через анализ сжатых признаков, после чего на выходе этапа классификации событие классифицируется как нормальное или аномальное.

Метод, предложенный Дэном Сюемом, имеет следующие преимущества:

- Устойчивость к шуму благодаря разреженным автоэнкодерам, которые подавляют помехи, сохраняя семантику движений.
- Анализ только динамических изменений используя двойное слияние для отделения и игнорирования фона.

К недостаткам можно причислить:

- Сложная настройка автоэнкодера, для которой требуется экспертиза по выбору оптимальной архитектуры.
- Риск переобучения при малом объеме данных - автоэнкодер запоминает шумовые паттерны.

Юань Гао [39] предлагает метод под названием Oriented Violent Flow (OVIF), который фокусируется на статистическом анализе изменений движения. Этот метод улучшает идентификацию насильственных действий путем использования линейного SVM-классификатора и адаптивных алгоритмов, таких как AdaBoost. Метод, предложенный Юанем Гао включает в себя три этапа. Этапы работы показаны на рисунке 8

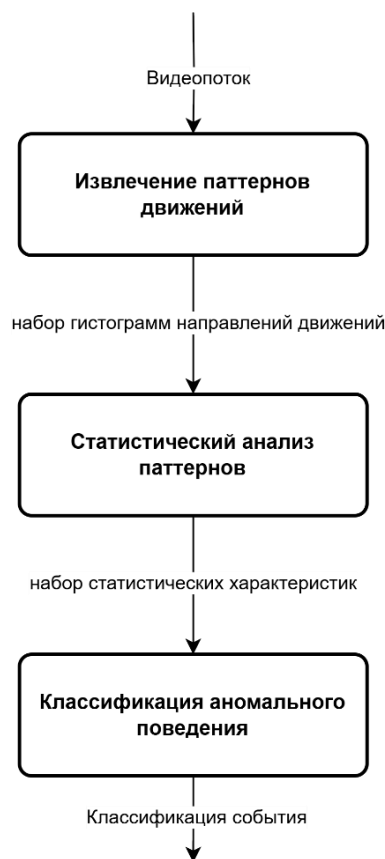


Рисунок 8 – Метод Юаня Гао

На первом этапе извлекаются признаки, связанные с резкими и интенсивными движениями, типичными для насилия с помощью анализа оптического потока и построения гистограммы направлений. В результате чего видеопоток преобразуется в набор гистограмм направлений движений, которые служат основой для дальнейшего анализа.

Второй этап преследует цель определения паттернов движения, которые указывают на насилие. Во время анализа паттернов для каждой гистограммы рассчитываются следующие параметры:

- Среднее значение интенсивности движений.
- Дисперсия направлений, то есть характеристика, показывающая величину хаотичности движений.
- Асимметрия распределения движений, то есть преобладание движений в одну сторону.

Далее на основе исторических данных строится эталонное распределение статистик для нормальных сцен, например, равномерное перемещение людей. Эталонное распределение статистик передается на следующий этап, где происходит классификация поведения в кадре.

На этапе классификации используется ансамбль из линейного SVM и AdaBoost. Линейный SVM обучается на полученных статистических признаках, для разделения классов поведения на аномальные и нормальные, то есть если значения признаков отличаются от эталонных на значение превышающее допустимое отклонение, то событие классифицируется как аномальное. AdaBoost усиливает веса сложных примеров, на которых SVM может ошибаться, например, сцены с частичным перекрытием объектов. Для адаптации системы к условиям пороги чувствительности SVM корректируются в зависимости от плотности толпы, освещения и других факторов.

Метод OViF обладает следующими преимуществами:

- Высокая точность в специфических сценариях благодаря фокусу на статистике оптического потока.
- Адаптивность к условиям, которая обеспечивается автоматической настройке параметров SVM и AdaBoost.
- Скорость обработки данных – использование линейного алгоритма SVM позволяет обрабатывать кадры в реальном времени.

- Устойчивость к фоновым помехам за счет статистического анализа гистограмм.

Недостатками метода являются:

- Специализация на одном типе аномальных событий - метод не адаптируется для обнаружения аномальных событий разного типа.
- Требуются размеченные данные для обучения SVM и AdaBoost.
- Неэффективность в медленных сценах, то есть в условиях низкой динамичности движений метод может не реагировать на событие, считая поведение нормальным.

Важным вкладом является работа J.F.P. Koopja [40], в которой представлена система под названием CASSANDRA для выявления враждебного поведения в общественных местах. Эта система объединяет аудио- и видеопотоки, применяя глубокие нейронные сети для прогнозирования поведения. Для достижения высокой точности системе требуется большой набор обучающих данных, охватывающий разнообразные сценарии взаимодействий. Интеграция мультимодальных данных позволяет CASSANDRA учитывать не только визуальные признаки, но и звуковые сигналы, такие как крики или шум, что повышает надежность обнаружения аномалий. В этом алгоритме этапы извлечения признаков, анализа данных и классификации событий организованы в единый конвейер. Система демонстрирует потенциал для снижения ложных срабатываний за счет корреляции между аудио- и видеоканалами, обеспечивая более точное распознавание агрессивных действий в условиях многолюдности. Однако реализация CASSANDRA сталкивается с вызовами, связанными с необходимостью синхронизации мультимодальных данных и высокими требованиями к вычислительным ресурсам, что ограничивает ее применимость в бюджетных устройствах. Несмотря на это, подход Koopja открывает перспективы для создания более устойчивых моделей, способных адаптироваться к изменяющимся условиям окружающей среды и обеспечивать

комплексный анализ поведения в общественных пространствах. Общее описание алгоритма CASSANDRA представлено на рисунке 10



Рисунок 10 – Алгоритм CASSANDRA

Этап извлечения признаков отличается от вышеперечисленных методов, тем что работает не только с видеоданными, но и со звуковой дорожкой. Для извлечения иерархических признаков с кадра используется VGGNET, который извлекает иерархические признаки как края и формы объектов, а также контекст окружающий объект. Аудиоданные перед анализом преобразуются в спектрограммы и текстовые метки через распознавание речи, то есть крики, повышенный тон голоса и резкие звуки выделяются как потенциальные признаки агрессии.

На втором этапе интеграции мультимодальных данных выполняется задача объединения видеопризнаков и аудиопризнаков в единую модель для анализа. Векторы из VGGNet (для видео) и спектограмма (для аудио) объединяются в единый набор данных, который синхронизируется по временным меткам, для связи визуальных событий со звуками.

Этап моделирования нормальных паттернов основывается на статистических данных, полученных с записей с нормальным поведением. Для каждого сценария формируются гистограммы визуальных и звуковых признаков. В результате моделирования нормальных паттернов формируются эталонные модели, которые используются для сравнения с текущими данными

Обнаружение аномалий происходит с использованием глубоких нейронных сетей, которые анализируют последовательности мультимодальных данных для выявления сложных сценариев. После обнаружения аномалия проходит процесс классификации с помощью SVM.

Метод выгодно отличается от выше предложенных по нескольким аспектам:

- Повышенная точность обнаружения аномальных событий за счет сочетания видеоанализа и звукового анализа.
- Устойчивость к фоновым помехам за счет кросс-проверки.
- Возможность реагирования на событие вне кадра, опираясь только на аудиоанализ.

К основным недостаткам можно отнести:

- Требование большого объема размеченных данных, как видео, так и аудио для обучения моделей.
- Чувствительность к качеству оборудования.
- Повышенная вычислительная нагрузка, которая затрудняет работу в реальном времени на устройствах, не оснащенных высокопроизводительными графическими процессорами.

Синьлу Цзун [41] разрабатывает метод, ориентированный на разреженное представление, в котором используются как обычные, так и аномальные видеоданные для создания моделей, классифицирующих поведение. Схематическое описание метода показано на рисунке 11

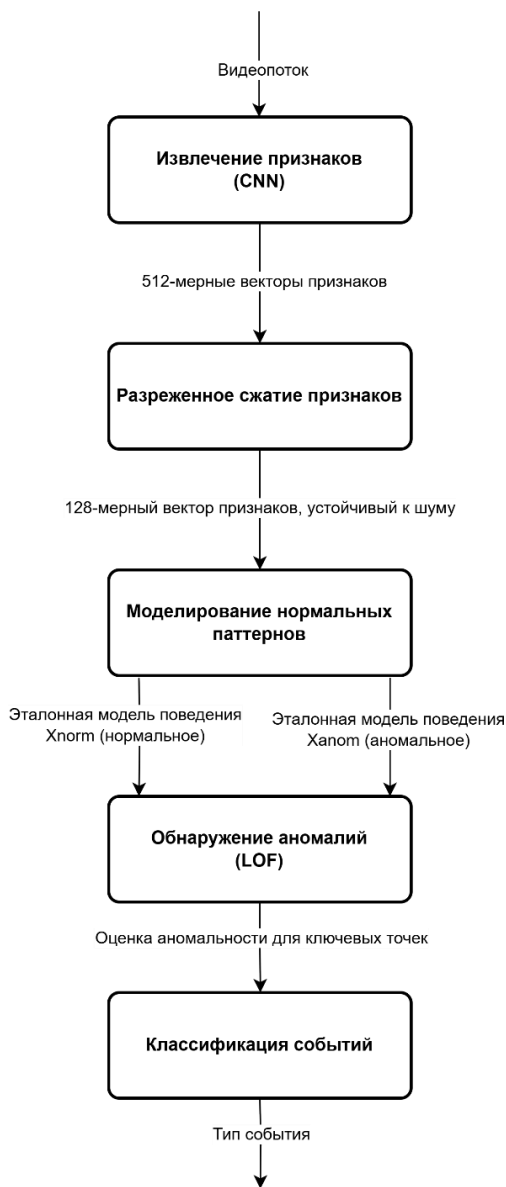


Рисунок 11 – Метод Синьлу Цзуна

На первом этапе признаки извлекаются с кадров с применением сверточных нейронных сетей (CNN). Также как в некоторых вышеописанных методах, извлекаются признаки трех уровней:

- Низкоуровневые - края, текстуры, градиенты.

- Средние уровни - Формы объектов.
- Высокоуровневые - Контекст сцены, то есть движение толпы или одиночные действия.

В результате чего видеопоток преобразуется в 512-мерные векторы признаков

После этапа извлечения признаков 512-мерный вектор передается на второй этап – разреженное сжатие, где посредством оптимизации он преобразуется в 128 мерный вектор признаков, при этом 128 мерный формат позволяет уменьшить влияние шумов и избытка данных.

На третьем этапе создаются эталонные модели для паттернов, которые будут определены как норма и паттернов аномального поведения. Модель проходит обучение на объединенном наборе данных, используя GMM для разделения признаков. Для каждого кадра строится вероятностная карта, описывающая вероятность принадлежности к норме.

Заключительным этапом является – обнаружение аномалии, цель которого состоит в идентификации отклонения от эталонных паттернов. На данном этапе используется Local Outlier Factor (LOF) который рассчитывает локальную плотность точек в пространстве признаков и принимает решение об аномальности события на основе отклонения, рассчитанного LOF. После обнаружения аномальное событие классифицируется ансамблевым голосованием SVM + AdaBoost.

Из основных плюсов данного подхода можно выделить следующие:

- Устойчивость к шуму - разреженное представление подавляет фоновые помехи.
- Возможность обнаружения аномалий различных типов.
- Снижение риска ложного срабатывания - комбинация LOF и статистического детектора снижает ложные срабатывания.

Но метод не лишен недостатков, таких как зависимость от качества видеопотока, сложность выборки данных для обучения, заключающуюся в том, что для обучения требуются как "нормальные", так и "аномальные" видеозаписи

Исследования в этой области продолжают уделять внимание двум основным аспектам: выделению различных признаков и разработке классификационной модели, что помогает улучшить точность и эффективность систем обнаружения аномальной активности.

Извлечение признаков — это ключевой этап в процессе анализа видеопотока, направленный на получение выделяющихся характеристик, указывающих на аномальные события. Эффективность обнаружения аномалий тесно связана с качеством извлеченных признаков, что подчеркивает важность выбора соответствующих методов и алгоритмов. Важные подходы, применяемые для извлечения признаков, включают гистограмму градиентов (HOG), архитектуры глубокого обучения, такие как VGGNet, а также метод переноса обучения.

Гистограмма градиента направления (HOG) представляет собой метод, который собирает информацию о направленности градиента в локальной области изображения для формирования гистограммы, отражающей контуры объектов в видеопотоке. Этот метод доказал свою эффективность в обнаружении краев и форм объектов, важных для анализа сцены.

Перед использованием HOG для извлечения признаков, в этой работе осуществляется бинаризация области переднего плана, что позволяет выделить объекты, представляющие интерес. Поскольку большинство пикселей в бинаризованном изображении имеют значение 0, это существенно сокращает время, необходимое для извлечения объектов. Таким образом, метод HOG оптимизирует процесс анализа, фокусируясь на значимых признаках, что улучшает общую эффективность обнаружения аномалий.

VGGNet, разработанная командой Оксфордского университета под руководством Карена Симоняна и Эндрю Зиссермана, является одной из самых

известных архитектур глубокого обучения и заняла второе место на конкурсе ILSVRC 2014. Архитектура VGGNet отличается своей простотой и состоит из последовательных сверточных слоев, за которыми следуют слои объединения. В зависимости от конкретной версии (VGG16 или VGG19) сеть включает в себя 16 или 19 сверточных слоев и плотную нейронную сеть с двумя скрытыми слоями и выходным слоем. Преимущество VGGNet заключается в ее способности извлекать сложные визуальные признаки, благодаря чему она может быть эффективно адаптирована для анализа видео. Эта модель особенно полезна для классификации объектов и их локализации, что делает ее полезным инструментом в системах обнаружения аномалий.

Метод трансферного обучения представляет собой стратегию, позволяющую ускорить процесс обучения новой нейронной сети, основываясь на знаниях, полученном от уже обученной модели. Например, если имеется обученная сеть для распознавания лиц, можно использовать ее базовые слои для обучения другой сети, предназначенной для распознавания причесок. В этом случае контрольные параметры и весовые коэффициенты нижних слоев первой сети могут быть перенесены в новую сеть, что позволит избежать необходимости обучать все низкоуровневые структуры заново. Сравнение методов обнаружения аномальной активности человека, использующих подход машинного обучения приведено в таблице 2

Таблица 2 - Сравнение методов обнаружения аномальной активности

Метод	Автор	Преимущества	Недостатки
Графы событий + нечеткая логика	Динеш Сингх	Моделирование взаимодействий между объектами, устойчивость к неопределенным данным, гибкость настройки под сценарии	Высокая вычислительная сложность, зависимость от качества трекинга и построения графа

Продолжение таблицы 2 - Сравнение методов обнаружения аномальной активности

GMM + оптический поток	Фейпинг Ли	Обнаружение аномалий без разметки, устойчивость к шуму, моделирование многомодальных распределений	Чувствителен к перекрытиям в толпе, требует ручной настройки порога нормального поведения
GMM + SVM	Чжаохуэй Ло	Устойчивость к фоновым помехам, автоматическая адаптация порогов, работа без разметки	Неэффективен в сложных сценах из-за упрощенных статистических моделей
OViF + AdaBoost	Юань Гао	Высокая точность в специфических сценариях, адаптивность к условиям, скорость обработки в реальном времени	Специализация только на насилии, необходимость разметки, чувствителен к качеству видеопотока
Пространственно-временной анализ (STIP + MoSIFT)	Де Соуза, Ниивас	Не требует трекинга, совместим с немаркированными данными, применим при средней плотности толпы	Чувствителен к шуму и климатическим условиям, низкая точность в динамичных сценах
Autoencoder + SVM / AdaBoost	Дэн Сюй	Фильтрация шума, снижение размерности данных до 128 бит, возможность работы с немаркированными данными	Сложная настройка автоэнкодера, риск переобучения при малом объеме данных
CASSANDRA (видео + аудиоанализ)	Пепайн Койжа	Повышенная точность за счет мультимодального анализа, устойчивость к фоновым помехам, реакция вне кадра через звук	Требуется большое количество размеченных данных, высокие требования к оборудованию
FFT + теорема Парсеваля	Илан Ван	Высокая чувствительность к резким изменениям, работа без разметки, устойчивость к фоновым изменениям	Не классифицирует типы событий, требует эталонного моделирования частотных паттернов
ST-Motion Patterns + GMM	Чжаохуэй Ло	Автоматическое обнаружение аномалий, снижение влияния шума, интеграция с SVM для классификации	Ограниченная применимость в условиях плотной толпы, не учитывает пространственные координаты аномалий
SVM + AdaBoost	Юань Гао	Высокая точность в специфических сценариях, адаптивность параметров, реальное время обработки	Специализация только на насилии, необходимость наличия как нормальных, так

Существующие подходы к обнаружению аномалий в видеопотоках, такие как использование пространственно-временных данных, методы машинного обучения (включая глубинное обучение) и продвинутое техники извлечения признаков продемонстрировали свою эффективность [44-45]. Однако они часто сталкиваются с ограничениями в динамичных и сложных сценах. В условиях высокой плотности объектов, непрерывного изменения фона и множества взаимодействий между объектами точность обнаружения может значительно снижаться.

1.3 Применение комбинированного подхода для обнаружения аномальной активности человека

Для улучшения результатов в динамичных сценах предлагается комбинированный подход, сочетающий несколько методов, каждый из которых вносит свой вклад в конечный результат. Предполагаемая многоуровневая система обнаружения аномалий представляет собой подход, который объединяет сильные стороны существующих методов и интегрирует их элементы для повышения точности и адаптивности анализа в динамичных сценах.

Архитектура системы включает следующие ключевые компоненты:

- Предобработка и извлечение признаков с помощью трансферного обучения с VGGNet [46] и автоэнкодерами. Где для извлечения признаков видеоизображений используется предобученная архитектура VGGNet, что минимизирует потребность в большом объеме данных. Автоэнкодеры с разреженными слоями применяются для выделения наиболее информативных характеристик и подавления шумовых составляющих.
- Пространственно-временной анализ с использованием оптического потока и быстрого преобразования Фурье (FFT) [47]. Оптический поток позволяет отслеживать движение объектов в реальном времени. Данные

преобразуются с помощью FFT для анализа частотных паттернов, что позволяет выявлять аномалии в динамике без ручной разметки.

- Моделирование нормальных паттернов с помощью гауссовских смесей (GMM) [48] и графов событий. GMM создает вероятностные карты нормального поведения, тогда как графы событий фиксируют связи между объектами на основе оптического потока. Такие графы строятся только для ключевых точек интереса, что снижает вычислительную нагрузку.

- Обнаружение аномалий с использованием Local Outlier Factor (LOF) и статистического детектора. LOF помогает выявлять локальные отклонения в данных. Статистический детектор проверяет эти отклонения на соответствие заранее определенным гипотезам, таким как всплески активности, что делает систему более устойчивой к шуму.

- Классификация и постобработка с помощью ансамбля методов SVM и AdaBoost. Метод опорных векторов (SVM) используется для классификации аномалий по выявленным паттернам. AdaBoost повышает точность классификации, адаптируясь к сложным сценариям. Окончательное решение принимается на основе голосования ансамбля, что позволяет снизить риск переобучения.

Представленные методы взаимодействуют между собой, дополняя друг друга. Каждый этап системы передает результаты следующему, создавая конвейер обработки видеопотока, где каждый компонент решает свою задачу. Такой подход позволит разработать универсальный инструмент, который может адаптироваться к различным типам аномалий, включая насилие, кражи и падения. Также повысить устойчивость к шуму, так как сочетание методов LOF и статистических детекторов делает систему менее чувствительной к фоновым помехам. Система может работать в реальном времени, благодаря оптимизированному конвейеру обработки который поддерживает высокую частоту кадров, что важно для систем, работающих в условиях реального времени.

Новизна подхода обуславливается использованием комбинацией методов:

- Симбиоз FFT и оптического потока, в отличие от методов, которые используют только временные ряды или статические признаки, то есть модель анализирует аномалии в частотной области. Это повышает чувствительность к резким изменениям активности, например, внезапным движениям.
- Интеграция GMM с графами событий позволяет моделировать не только локальные паттерны, но и глобальные связи между объектами, устраняя ограничения, возникающие при перекрытии объектов.
- Совместное использование SVM и AdaBoost сочетает сильные стороны SVM (высокая точность на линейно разделимых данных) и AdaBoost (адаптивность к сложным распределениям), превосходя методы, основанные на едином классификаторе.
- Использование разреженных автоэнкодеров в совокупности с трансферным обучением сокращает размерность данных, сохраняя информативность признаков.

Предполагается, что метод сможет адаптироваться к различным сценариям за счет гибкого моделирования нормальных паттернов и мультимодального анализа. Это отличает его от узкоспециализированных методов, таких как OViF.

1.4 Постановка задачи управления системой

Задача управления в рамках разработки интеллектуальной системы видеонаблюдения заключается в обеспечении автоматизированного мониторинга и анализа видеопотока с целью выявления аномального поведения человека в реальном времени.

Модель обнаружения аномальной активности является частью основной системы видеонаблюдения для обнаружения аномальной активности человека, эта система разделена включает в себя пять подсистем, которые выполняют задачи по обработке изображения.

- Подсистема предобработки и извлечения признаков (P_i) содержит в себе модуль нормализации MD и подавитель шума ND, которые обрабатывают входной кадр $F_{\text{кадр}}$.
- Подсистема пространственно-временного анализа (A_i) извлекает из полученных кадров оптический поток OF и преобразует данные в частотную область помощью модуля преобразования данных MP.
- Подсистема моделирования нормальных паттернов (M_i) работает с вероятностными картами PM и графами событий GS, применяя фильтр точек интереса FI к значениям, которые могут указывать на аномальные события.
- Подсистема обнаружения аномалий (D_i) принимает вероятностные карты PM и графы событий GS от прошлой подсистемы, а также частотные данные QD, полученные из модуля преобразования данных на втором этапе.
- Подсистема классификации и постобработки содержит в себе классификацию события по типу аномалии CS, а также координаты обнаруженного события XY, для отображения на кадре.

Организационной системой выступает любое учреждение, заинтересованное в обеспечении безопасности через использование систем видеонаблюдения, например общественные места в городе или помещения с высокой посещаемостью. Из задач надсистемы, которые планируется решить с использованием системы видеонаблюдения для обнаружения аномальной активности человека, можно выделить: обеспечение безопасности на подконтрольной территории, минимизация рисков от преступлений, террористических актов и других угроз, снижение зависимости от человеческого фактора при мониторинге. При этом внедрение системы позволит оргсистеме снизить затраты на мониторинг систем видеонаблюдения и увеличить скорость реакции на инциденты.

Внешняя среда включает в себя следующие элементы:

- Условия съемки, такие как освещение, погодные условия, фоновая активность.

- Социальная среда, то есть характеристики объектов наблюдения - плотность толпы, взаимодействие между людьми [49].
- Технические факторы качество видеопотока, состояние оборудования.

Входными данными для оргсистемы является запрос на анализ видеопотока, для системы - видеопоток с камер наблюдения, который является внутренним потоком информации. Выходными данными для системы являются классификация аномальности события, визуализация аномалий на видео.

Процессы, происходящие в системе представлены следующими подсистемами:

- Подсистема предобработки и извлечения признаков (P_i) отвечает за подготовку данных для анализа, т.е. получает на вход видеопоток с камеры $F_{\text{кадр}}$ и после обработки с помощью модулей MN - модуля нормализации и ND - подавителя шума предоставляет нормализованные и очищенные данные в виде 128-мерных векторы признаков x_{norm} для каждого кадра. Подсистема включает в себя
 - Подсистема пространственно-временного анализа (A_i) выполняет анализ динамики сцены для выявления аномальных движений, получая на вход векторы признаков x_{norm} , полученные на предыдущем этапе и преобразовывает их в частотные паттерны E_f с очищенными векторами оптического потока $v=(u,v)$.
 - Подсистема моделирования нормальных паттернов (M_i) создает эталонные модели "нормального" поведения для последующего сравнения, принимая на вход выходные данные с подсистемы пространственно-временного анализа, преобразуя их в вероятностные карты $p(x)$, графовые паттерны и графовые паттерны взаимодействий.
 - Подсистема обнаружения аномалий (D_i) идентифицирует отклонения от нормальных паттернов. Входными данными для неё являются вероятностные карты $p(x)$, графовые паттерны, частотные паттерны $E(f)$, выходными данными - оценки аномальности ($LOF(x)$, p -value) и координаты ключевых точек.

- Подсистема классификации и постобработки (C_i) классифицирует события как "аномалия" или "норма" и готовит результаты для вывода, то есть принимает данные от подсистемы обнаружения аномалий и выдает классификацию аномальности события и визуализацию области с аномальным событием.

Взаимодействие между подсистемами описывается как $V(t) = P_i + A_i + M_i + D_i + C_i$

Надсистема и система взаимодействуют через передачу данных видеопотока и результатов анализа. В рамках надсистемы инновационный инструмент обработки видеоданных используется для автоматизации процесса обнаружения аномалий. Полученные данные (например, координаты аномалий, типы событий) затем передаются в надсистему для дальнейшей обработки и принятия решений

Эффективность системы зависит от следующих параметров:

- $t_{откл.}$ - время отклика системы, то есть время между получением кадра и формированием сигнала об аномалии.
- $A_{точн.}$ - точность обнаружения аномального поведения.
- $F_{лс.}$ - частота ложных срабатываний
- $O_{кол.}$ - максимальное количество одновременно обрабатываемых объектов
- $G_{дост.}$ - доступность системы
- $P_{рес.}$ - потребление ресурсов

Формулу эффективности системы можно представить в виде:

$$A = \frac{A_{точн.} * O_{кол.}}{t * F_{лс.} * P_{рес.}} * G_{дост.}$$

Надсистема осуществляет управление системой с помощью следующих векторов управления:

- U_{funct} - управление функционированием системы, то есть изменение порогов чувствительности.
- U_{res} - управление составляющими системы, например замена видеокамер на более качественные, для повышения качества обработки, или увеличение мощностей обработки для более быстрой работы системы.

Схема управления системой представлена на рисунке 12

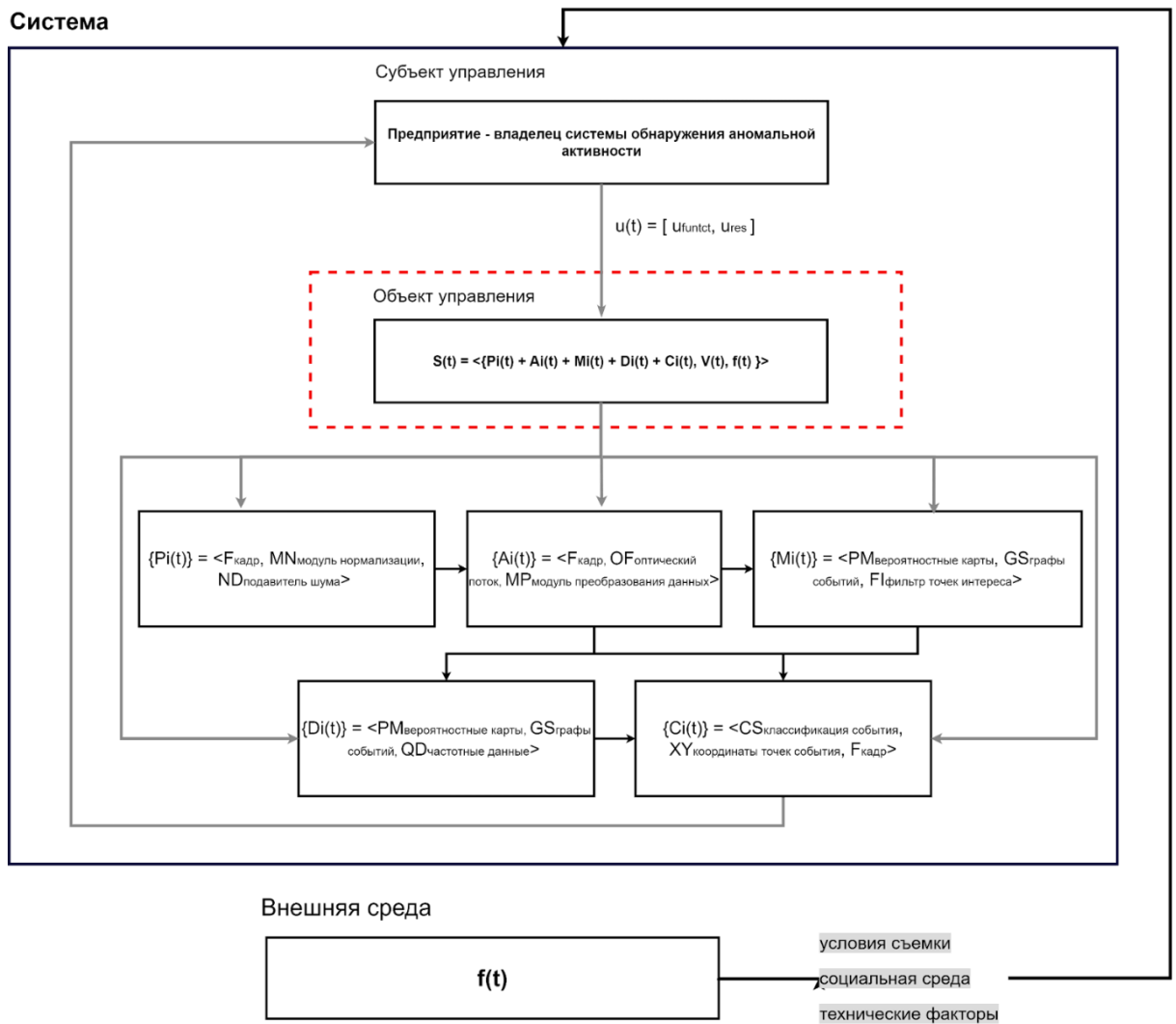


Рисунок 12 – Постановка задачи управления системой

1.5 Результаты и выводы первой главы

В первой главе был проведен анализ существующих подходов, сформирована актуальность исследования, предложено решение задачи

обнаружения аномальной активности человека с помощью многомодульной интеллектуальной системы, а также раскрыта задача управления системой.

Результаты анализа в первом разделе показывают, что традиционные системы видеонаблюдения в современных условиях сталкиваются с рядом проблем, из-за которых эффективность обнаружения снижается, например:

- Проблемы, возникающие из-за современных тенденций.
- Проблемы влияния человеческого фактора.
- Технические проблемы.

Таким образом была обусловлена актуальность разработки гибридной модели, способной работать с немаркированными данными, устойчивой к шуму и адаптивной к внешним факторам и в итоге определять аномальное поведение человека в кадре в автономном режиме, без участия оператора.

Во втором разделе были проанализированы существующие научные подходы для обнаружения аномального поведения в потоке видео. В каждом научном подходе выделены методы их достоинства и недостатки. Список проанализированных подходов выглядит следующим образом:

- Подходы моделирования социологических моделей.
- Моделирование когнитивных процессов, которые формализуют эвристики и поведенческие модели.
- Гибридный подход Мохаммади который внедряет набор простых поведенческих эвристик в физические уравнения.
- Подходы, не требующие трекинга объектов, например.
- Подходы, основанные на пространственно-временных данных.
- Подходы, использующие алгоритмы машинного обучения.

В результате сравнения, для реализации был выбран подход, использующий алгоритмы машинного обучения, который является самым современным и совершенным из рассмотренных.

В третьем разделе первой главы было предложено решение проблемы

эффективности видеомониторинга с помощью создания комбинированной модели. Описана архитектура системы, а также способ взаимодействия модулей внутри системы.

На четвертом этапе была сформирована схема управления системой, описаны векторы воздействия на систему, внешняя среда и определена цель системы, которая направлена на повышение эффективности системы.

2 РАЗРАБОТКА МОДЕЛИ ОБНАРУЖЕНИЯ АНОМАЛЬНОЙ АКТИВНОСТИ ЧЕЛОВЕКА С ИСПОЛЬЗОВАНИЕМ МНОГОУРОВНЕВОВОГО ПОДХОДА

2.1 Описание разрабатываемой модели для обнаружения аномальной активности человека

Разрабатываемая модель должна решать задачу обнаружения аномальной активности в динамичных сценах, где традиционные методы сталкиваются с ограничениями, такими как: шум, перекрытие объектов, высокая вычислительная нагрузка. Для решения данной задачи предлагается использовать комбинированную структуру, объединяющую современные подходы машинного обучения и анализа данных. Модель, содержащая в себе модули преобразования и анализа данных представлена на рисунке 13

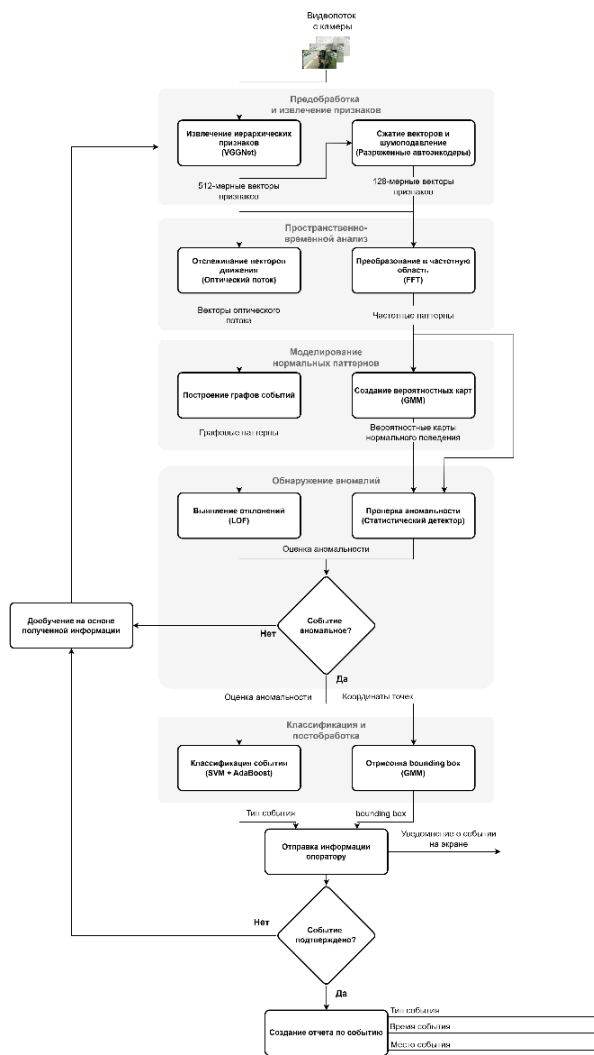


Рисунок 13 – Схема разрабатываемой модели

2.2 Анализ существующих моделей для предобработки изображений и извлечения признаков

Задача извлечения признаков состоит в улавливании паттернов изображения, например движения людей и их взаимодействие. Выбранная модель VGG16 содержит 13 сверточных слоев, которые извлекают иерархические признаки трех уровней: Низкоуровневые (текстуры, края), признаки среднего уровня (формы объектов) и высокоуровневые (контекст сцены).

Аналог в виде ResNet [50] более эффективен для классификации, однако остаточные связи данной требуют больше вычислительных ресурсов, по сравнению с VGG16, что является нежелательной характеристикой для обработки в режиме реального времени. С другой стороны, модель можно сравнить с MobileNet [51], которая превосходит VGG16 в скорости работы, но при этом обеспечивает меньшую точность обнаружения аномалий в сложных сценах. Минимизировать зависимость от размеченных данных позволяет трансферное обучение, то есть использование уже предобученной на ImageNet модели.

Таким образом модель VGG16 предобученная на наборе ImageNet является сбалансированным решением, для этапа извлечения признаков с входного изображения.

С целью выделения наиболее информативных паттернов используются автоэнкодеры, которые сжимают и очищают извлеченные признаки от шумов и устраняют избыточность данных. Энкодеры преобразуют входные признаки, полученные от модели VGG16 в сжатое представление, что ускоряет обработку данных на последующих этапах. Свойство разреженности фокусирует модель на ключевых характеристиках, то есть на движениях людей и игнорирует посторонний шум, такой как тени и изменение изображения из-за осадков или других погодных явлений.

В сравнении с методом главных компонент PCA [52], который также используется для сжатия признаков, автоэнкодеры являются нелинейным методом, то есть более эффективны для нелинейных данных, таких как видеопоток, также PCA сохраняет вариативность данных, а не их семантику, то есть использование метода главных компонент невозможно для подавления шумов. Более сложные вариационные автоэнкодеры, по сравнению с используемыми разреженными требуют больших объемов данных и более ресурсоемки из-за использования вероятностного распределения для каждого признака, когда как в разреженных признаки хранятся как дискретные значения в слоях признаков.

Итого для первого этапа работы алгоритма, который заключается в извлечении и сжатии признаков для дальнейшего анализа были выбраны модели VGG16 с применением разреженных автоэнкодеров. После завершения первого этапа алгоритм получает на выходе подготовленные данные с минимальными потерями информации и высокой устойчивостью к помехам.

На втором этапе работы алгоритма проводится анализ извлеченных признаков, для выявления аномальных паттернов движения во временной и частотной областях. В процессе выполнения этапа между кадрами отслеживаются векторы движения объектов, затем временные ряды преобразуются в частотную область, для выявления аномальных колебаний.

Отслеживание движения объектов осуществляется методом оптического потока, используя алгоритм Лукаса-Канаде, который вычисляет векторы движения для ключевых точек, то есть алгоритм анализирует смещение пикселей на последующих кадрах, тем самым определяя движение обнаруженного объекта. Также для отслеживания можно использовать Farneback [53], имеющий хорошие результаты в отслеживании плавных движений, но алгоритм Farneback не эффективен для отслеживания резких движений. Методы глубокого обучения, такие как FlowNet, требуют больших вычислительных мощностей по сравнению с алгоритмом оптического потока Лукаса-Канаде, который имеет линейную

оценку сложности, что позволяет проводить отслеживание в реальном времени, не требуя больших вычислительных мощностей.

Далее для выявления аномальных колебаний требуется преобразовать векторы оптического потока в частотную область, применяя быстрое преобразование Фурье (FFT), это позволит отделять области с резкими движениями, обнаруживая высокие частотные показатели, без необходимости ручной разметки. Альтернативами подхода FFT являются временные ряды, такие как RNN [54] или LSTM [55], для работы которых требуется обучения на размеченных данных, также такой подход хуже работает с кратковременными аномалиями. Статистические методы, например вычисление среднего значения, имеют меньшую точность по сравнению с частотными характеристиками.

Третий представляет собой моделирование нормальных паттернов, состоящие из вероятностных распределений движений и графов событий. В процессе выполнения этого этапа создаются эталонные модели нормального поведения для последующего сравнения с текущими событиями. Это позволяет системе различать типичные действия такие как движение толпы и аномалии, например насилие.

Моделирование вероятностных распределений нормальных движений на основе частотных паттернов, полученных на предыдущем этапе, происходит с использованием модели Гауссовской смеси (GMM), механизм работы которой оценивает параметры гауссовских компонентов и рассчитывает карты вероятностей для каждого кадра, где низкие значения указывают на аномальное событие. В отличие от одиночных гауссиан GMM имеет возможность моделирования многомодальных распределений, таких как разные типы движений в толпе, также GMM устойчива к шуму благодаря вероятностному подходу. Сравнивая GMM с моделями автоматической регрессии, можно выделить, что такие модели имеют меньшую эффективность для пространственных данных.

Построение графов на основе ключевых точек оптического потока, служит для анализа взаимодействий между объектами. Данными для построения служат векторы оптического потока с предыдущего этапа работы модели. При построении выделяются ключевые точки, т.е. зоны с высокой динамикой, после чего граф разбивается на кластеры для выявления типичных взаимодействий. Графы учитывают глобальные связи недоступные для методов, анализирующих отдельные объекты, а также снижают вычислительную нагрузку за счет фокусировки на ключевых точках, что в сравнении с альтернативными методами, такими как трекинг объектов, который требует сложной идентификации каждого объекта и не является оптимальным для динамичных сцен, выделяет метод построения графов для анализа взаимодействий объектов. Традиционные методы статистики, например вычисление средней скорости игнорируют структурные взаимодействия, из-за чего существенно снижается значимость информации в сценариях со сложными сценами с несколькими объектами.

В результате выполнения данного этапа, объединяя статистические и структурные методы, алгоритм формирует эталоны для обнаружения аномалий в виде вероятностных карт и графовых паттернов.

Следующий этап - обнаружение аномалий, задачей которого является идентификация отклонений от нормальных паттернов, выявленных на предыдущих этапах, с использованием комбинации Local Outlier Factor (LOF) и статистического детектора.

Local Outlier Factor (LOF) выявляет локальные отклонения в данных, которые могут быть неочевидны при глобальном анализе, то есть для каждой точки рассчитывается отношение локальной плотности её соседей к её собственной плотности. Такой механизм работы позволяет обнаруживать аномалии в сценах с неравномерной плотностью данных, например, резкое движение одного человека в толпе, по сравнению с методами кластеризации,

например, DBSCAN [56], LOF не требует ручной настройки параметров, из-за чего лучше адаптируется к динамичным сценам.

Статистический детектор полезен для проверки гипотез о наличии аномальных всплесков активности. Для каждого кадра проверяется гипотеза - «событие нормальное» против «событие аномальное». Статистический детектор автоматически адаптируется к изменениям в данных и не требует разметки, в отличие от бинарных классификаторов.

Выходными данными четвертого этапа являются оценки аномальности для каждого кадра и координаты ключевых точек с высоким LOF фактором.

Заключительным этапом работы алгоритма являются классификация и постобработка, на котором принимается окончательное решение об аномальности события. Входные данные в виде оценок аномальности и паттернов (частотных и графовых, полученных на этапе моделирования нормальных паттернов).

Классификация событий на основе линейно разделимых признаков выполняется с помощью ансамбля SVM (Support Vector Machine) + AdaBoost (Adaptive Boosting). SVM использует RBF-ядро для разделения линейно неразделимых данных, что повышает точность распознавания аномального события. AdaBoost последовательно усиливает веса сложных для классификации примеров, что улучшает адаптивность к сложным сценариям. Объединение методов в ансамблевое голосование снижает риск переобучения, превосходя методы с единичными классификаторами. Также подход с ансамблем SVM + AdaBoost требует меньших объемов данных по сравнению с глубокими нейронными сетями.

Постобработка служит для устранения ложных срабатываний и подготовки результатов для вывода. Первый шаг в постобработке - фильтрация изолированных аномалий, то есть удаление событий короче 0.5 секунд, чтобы отсеять случайные всплески. Далее происходит проверка локализации аномалий,

после чего отрисовывается bounding box [57] вокруг аномальных объектов. Постобработка повышает интерпретируемость результатов для операторов, а также снижает процент ложных срабатываний.

На выходе последнего этапа модель выдает классификацию события «аномальное» или «нормальное», а также bounding box, если событие было классифицировано как аномальное.

2.3 Выбор исходного набора данных для обучения модели

Набор данных для обучения модели (датасет) — это совокупность данных, подготовленных для обучения, валидации и тестирования модели. В контексте систем видеонаблюдения он включает видеозаписи, размеченные или немаркированные, с метаданными, описывающими сцену (например, тип помещения, плотность толпы, освещение). Датасет используется в моделях машинного обучения со следующими целями:

- Обучение модели, то есть размеченные данные в контексте исследования формируют знания модели о типичных и аномальных действиях.
- Тестирование и валидация – проверка точности обнаружений аномалий и устойчивости к помехам.
- Повышение адаптации к различным сценариям.

Для эффективного обучения и тестирования разрабатываемой системы обнаружения аномальной активности человека необходимо, чтобы набор данных соответствовал строгим критериям, обеспечивающим высокую точность, устойчивость к шуму и универсальность модели. Основные требования к набору данных выглядят следующим образом:

- Разнообразие сценариев аномального поведения: датасет должен охватывать разные типы аномальных событий.
- Разнообразие условий, то есть видеозаписи или изображения в наборе данных должны содержать различные социальные среды и условия съемки.

- Реалистичность сцен, чтобы модель сразу была адаптирована к реальным условиям.

- Объем данных не менее 10 тысяч файлов.

- Открытость датасета.

Из доступных датасетов были выбраны ShanghaiTech Campus [58], CUHK Avenue [59], UCF-Crime [60] и Violent-Flows [61]

ShanghaiTech Campus - разработанный для анализа поведения в многолюдных сценах, таких как университетские кампусы. Основная цель датасета — исследование массовых аномалий, включая давку, паническую реакцию и хаотичное движение в толпе. Данные сняты в реальных условиях с использованием камер видеонаблюдения, что делает его релевантным для задач обнаружения аномалий в общественных местах. Датасет содержит более 270 тысяч кадров и 130 аномальных событий, разделенных на 13 классов которым относятся такие аномалии как:

- Падения.

- Давка в толпе.

- Панические реакции.

- Нарушения структуры потока людей.

Записи в датасете содержат сцены с высокой плотностью людей, сцены сняты в различных условиях, таких как ночное время суток, дождь, снег. Видео в датасете имеют разрешение 256 на 256 пикселей с частотой кадров 25FPS. Метаданные включают плотность толпы и типы сцен. Пример данных находящихся в датасете представлен на рисунке 14



Рисунок 14 – Пример данных ShanghaiTech Campus

Набор данных имеет следующие преимущества:

- Высокое разрешение кадров (256×256), которое позволяет анализировать детали движения в условиях многолюдности.
- Реалистичность сцен - видеозаписи сняты в реальных условиях (университетский кампус), что повышает применимость модели к практическим задачам.
- Фокус на массовых событиях – датасет подходит для тестирования методов, ориентированных на толпу (например, графы событий, модели SFM).
- Различные погодные и световые условия, то есть набор включает сцены с дождем, снегом и низким освещением, что полезно для оценки устойчивости к фоновым помехам.
- Универсальность для исследований - широко используется в работах по анализу динамики толпы и предсказанию давки.

К недостаткам можно отнести:

- Отсутствие временной разметки аномалий - Затрудняет обучение моделей, требующих точных меток начала и окончания событий, например, SVM и LSTM.
- Ограниченная вариативность событий - специализируется на массовых аномалиях, не охватывает индивидуальные действия.

Следующий датасет - CUNHK Avenue, разработанный для анализа аномального поведения в динамичных сценах на кампусной улице. Он включает 16 обучающих и 21 тестовый видеоклип, собранный с камеры на территории CUNHK (Китайский университет Гонконга) . Общий объем данных составляет

30,652 кадра (15,328 в обучающих и 15,324 в тестовых видеозаписях). Датасет сопровождает работу Цей Лу [62], посвященную обнаружению аномалий с высокой частотой обработки (до 150 FPS). Данные содержат всего 3 класса аномальных событий такие как:

- Strange action: необычные действия, например человек бежит против потока.
- Wrong direction: движения в запрещенном направлении, например, пешеход на проезжей части.
- Abnormal object: появление неожиданных объектов, например, стоящий человек в толпе.

Пример данных представлен на рисунке 15



Рисунок 15 – Пример данных CUNK Avenue

К преимуществам относятся:

- Высокая частота кадров, которая позволяет анализировать быстрые движения.
- Реалистичность сцен – сцены сняты в реальных условиях (кампусный проспект), что делает их релевантными для задач мониторинга общественных мест.
- Пространственная разметка аномалий - прямоугольники (bounding boxes) позволяют точно локализовать источник аномалии, что важно для систем видеонаблюдения.
- Наличие аномальных объектов и направлений движения подчеркивает применимость датасета для обнаружения нетипичного поведения.

- Универсальность метрик оценки - использование критериев VOC Pascal обеспечивает объективное сравнение результатов с другими методами.

Из недостатков можно выделить:

- Ограниченное количество классов – датасет содержит всего 3 разметки, что делает его менее подходящим для многозадачного анализа.

- Отсутствие временной разметки - нет точных меток начала и окончания аномалий, что усложняет обучение моделей, требующих временной привязки (например, LSTM).

- Присутствует тряска камеры – в некоторых сценах камера не статична, что не соответствует реальному поведению камер наружного видеонаблюдения.

- Несбалансированность данные - В обучающих данных есть выбросы и редкие нормальные паттерны, что увеличивает риск переобучения.

Третий датасет, который рассматривается для обучения системы обнаружения аномальных событий - Violent-Flows, специализированный датасет для обнаружения насильственных действий в видеопотоках. Он был разработан в рамках работы Юань Гао. Датасет собран из YouTube-видео, что делает его источником реалистичных сцен, часто встречающихся в общественных местах. Основная цель датасета — тестирование алгоритмов, использующих пространственно-временные паттерны и ансамблевые методы. Набор содержит 246 видеофайлов, с продолжительностью от 1.04 до 6.52 секунд (в среднем 3.60 секунд). Классификация событий только на 2 типа, нормальное поведение или аномальное. Пример файлов из датасета показан на рисунке 16



Рисунок 16 – Пример данных Violent-Flows

Преимуществами набора данных Violent-Flows, являются:

- Поддержка бенчмарков, включает стандартные протоколы оценки, что позволяет объективно сравнивать методы.
- Наличие временной разметки, которая упрощает обучение SVM и AdaBoost.

К недостаткам:

- Ограниченное количество классов - фокус только на насилие и норму, без разделения на типы аномалий.
- Отсутствие пространственной разметки - нет точной локализации источника аномалии, например отметки с участниками драки.
- Короткие видеоклипы - средняя продолжительность 3.60 секунд затрудняет анализ сложных сцен, требующих контекста.
- Низкая вариативность условий - отсутствие данных о погоде, освещении и плотности толпы снижает универсальность модели.
- Видео не соответствует формату видеопотока с CCTV систем – т.е. видео снято с рук, что нетипично для камер видеонаблюдения.

Последний датасет - UCF-Crime — это масштабный датасет, разработанный для анализа разнообразных аномальных событий в видеопотоках. Он был собран на основе реальных записей с камер видеонаблюдения, включая 14 классов аномальных событий, таких как:

- Злоупотребление.
- Арест.
- Поджог.
- Нападение.
- Кража со взломом.
- Взрыв.
- Драка.
- Обычные видео.
- Дорожные происшествия.
- Ограбление.
- Стрельба.
- Магази́нная кража.
- Кража.
- Вандализм.

Датасет содержит кадры из записей CCTV камер, всего 264 видеоролика, из которых извлекались кадры с шагом в 10 секунд, для сокращения объема, также разрешение приведено к единому формату 64 на 64 пикселя. Датасет также содержит размеченные данные – временную разметку аномалий, то есть начало и окончание события. Пример данных из датасета представлен на рисунке 17

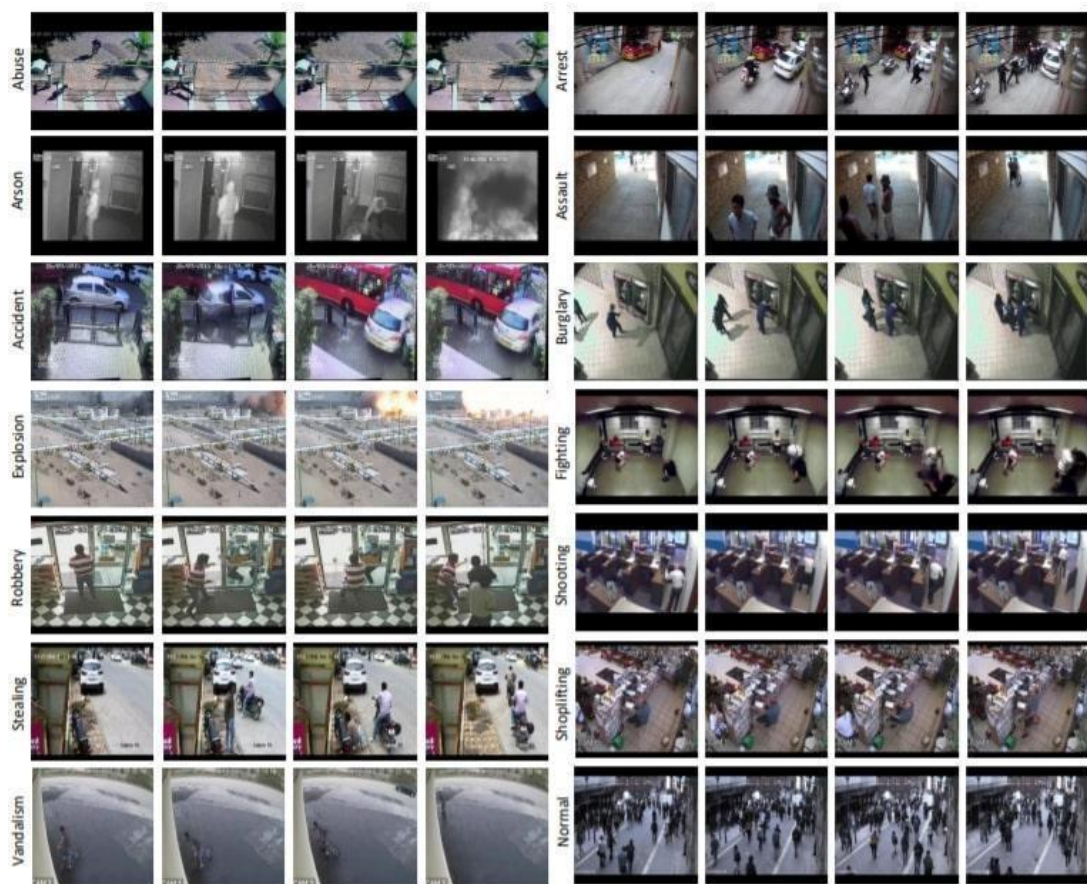


Рисунок 17 – Пример данных UCF-Crime

К положительным сторонам данного набора данных можно отнести:

- Реалистичность сцен - все записи сделаны с камер видеонаблюдения, что делает их релевантными для практического применения.
- Разнообразие аномалий - охватывает 14 классов событий.
- Поддержка временной разметки, которая позволяет обучать модели на точных моментах начала и окончания аномалий.
- Адаптивность к внешним условиям – включает сцены с различными погодными условиями и условиями освещения.

К недостаткам можно отнести:

- Низкое разрешение кадров.
- Дисбаланс классов, то есть требуется оверсемплинг или андерсемплинг.

Сравнение по основным параметрам представлено в таблице 3

Таблица 3 – Сравнение наборов данных по параметрам

Параметр	UCF-Crime	ShanghaiTech Campus	CUHK Avenue	Violent-Flows
Количество классов	14 (разнообразные)	13 (массовые события)	2 (норма / аномалия)	2 (насилие / норма)
Качество данных	Низкое разрешение	Высокое разрешение	Среднее разрешение	Среднее разрешение
Наличие разметки	Временная разметка (без пространственной)	Нет разметки	Временная и пространственная разметка (bounding boxes)	Только временная разметка
Устойчивость к шуму	Средняя (требует автоэнкодеров для фильтрации)	Высокая (оптимизирован для толпы)	Высокая	Средняя
Применимость к разным сценариям	Высокая (14 классов, индивидуальные и массовые аномалии)	Средняя (специализация на толпе)	Низкая (2 класса, узкая задача)	Низкая (только насильственные действия)
Требования к ресурсам	Умеренные (GMM, FFT)	Высокие (анализ толпы)	Низкие (кратковременные события)	Низкие (узкая задача)

При сравнении выявлены преимущества и недостатки каждого набора данных. Результаты сравнения датасетов для обучения модели с целью обнаружения аномального поведения человека, представлены в таблице 4

Таблица 4 – Сравнение наборов данных для обучения модели

Датасет	Плюсы	Минусы
UCF-Crime	14 классов, реалистичные сцены, временная разметка	Низкое разрешение, дисбаланс классов

ShanghaiTech Campus	Высокое разрешение, устойчивость к многолюдности	Ограниченные классы, отсутствие разметки
Campus Avenue	Высокая частота кадров, пространственная разметка	Узкая специализация, отсутствие мультимодальности
Violent-Flows	Точное обнаружение насилия, реальные сцены	Узкая задача, отсутствие пространственной разметки

Проведенный анализ четырех датасетов (UCF-Crime, ShanghaiTech Campus, Campus Avenue, Violent-Flows) показал, что UCF-Crime лучше всего соответствует требованиям разрабатываемой системы обнаружения аномальной активности человека. Его выбор обусловлен вышеописанными ключевыми преимуществами.

2.4 Результаты и выводы второй главы

Вторая глава - «Разработка модели обнаружения аномальной активности человека с использованием интеллектуальных систем видеонаблюдения» описывает разрабатываемую многомодульную систему. Работа системы включает в себя пять ключевых этапов:

1. Предобработка и извлечение признаков.
2. Пространственно-временной анализ.
3. Моделирование нормальных паттернов.
4. Обнаружение аномалий.
5. Классификация и постобработка.

В разделе с выбором моделей для каждого этапа обработки данных были выбраны следующие модели и инструменты:

- VGGNet + Разреженные автоэнкодеры для первого этапа.
- Оптический поток Лукаса-Канаде + Быстрое преобразование Фурье (FFT) для второго этапа.

- Построение графов событий на основе ядра SVM + вероятностные карты (GMM) для третьего этапа.
- Local Outlier Factor (LOF) + Статистический детектор для четвертого этапа.
- Ансамблевое голосование SVG и AdaBoost + Bounding box для пятого этапа.

В разделе, посвященном выбору исходного набора данных, проводится сравнительный анализ датасетов, определяются требования к набору данных. По результатам сравнения четырех датасетов:

- ShanghaiTech Campus.
- CUHK Avenue.
- UCF-Crime.
- Violent-Flows.

Было определено, что набор данных UCF-Crime является наиболее подходящим для обучения интеллектуальной системы.

3 ПРОГРАММНАЯ РЕАЛИЗАЦИЯ АЛГОРИТМА ОБРАБОТКИ ВИДЕОДАНЫХ ДЛЯ ВЫЯВЛЕНИЯ АНОМАЛЬНОЙ АКТИВНОСТИ ЧЕЛОВЕКА

3.1 Предпроектное обследование предприятия для внедрения алгоритма в системы видеонаблюдения

Внедрение автоматического мониторинга в реальном времени с использованием видеокамер может быть организовано с основной целью повышения уровня безопасности на предприятии или в общественном месте.

Рассмотрим текущее состояние процесса мониторинга аномальных событий без внедрения алгоритма автоматического обнаружения. Цель процесса - обеспечение безопасности в общественных местах. Участниками процесса являются:

- Оператор мониторинга - выполняет роль визуального анализа видеопотоков с камер в реальном времени.
- Службы безопасности - выполняют роль реагирования на обнаруженные инциденты.
- Техническая служба - выполняет роль обслуживания инфраструктуры видеонаблюдения, например установка и подключение новых камер или поддержка работоспособности сервера с данными.

Процесс мониторинга происходит непрерывно и содержит в себе пять этапов

На первом этапе сбора и анализа видеопотока данные собираются с помощью видеокамер и выводятся на монитор оператора в реальном времени. Анализ полученных данных выполняет оператор мониторинга. Проблемой данного метода является низкая эффективность из-за влияния человеческого фактора, так как оператор отслеживает события с нескольких камер, возникает когнитивная перегрузка, что влечет за собой снижение концентрации и увеличение времени реагирования [63]. Также возникает проблема

масштабируемости, то есть при увеличении количества камер необходимо наращивать штат операторов, что приводит к увеличению расходов.

На втором этапе обнаружения аномального события оператор обнаруживает событие и определяет его как аномальное или нормальное, посредством визуального контроля. Определение аномальности события полностью осуществляется оператором, в качестве вспомогательной системы может быть рассмотрена детекция движения, большинство систем видеонаблюдения имеют возможность реагирования на движения в кадре, это повышает вероятность обнаружения инцидентов на объектах, где необходима фиксация присутствия посторонних лиц, при этом зона наблюдения статична, например, закрытая территория или помещение. Для обеспечения безопасности в общественных местах простая система обнаружения движения бесполезна, так как динамичное движение в области видимости камеры является нормой.

Третий этап - реагирование на инцидент. Оператор уведомляет службы безопасности, дальнейшие действия осуществляются службами безопасности.

На четвертом этапе оператор фиксирует время, тип и место инцидента в отчетном журнале. Проблемой данного этапа является упущение контроля при заполнении журнала, так как внимание оператора не сфокусировано на видеопотоке.

В итоге основная проблема традиционного метода видеомониторинга - недостаточная эффективность отслеживания и обнаружения аномальных событий, которая возникает из-за влияния человеческого фактора. Также присутствует проблема масштабирования, которая подразумевает дополнительные расходы на расширение штата операторов [64] при увеличении зон покрытия видеокамер. Схема процесса мониторинга без использования интеллектуальной системы показана на рисунке 18

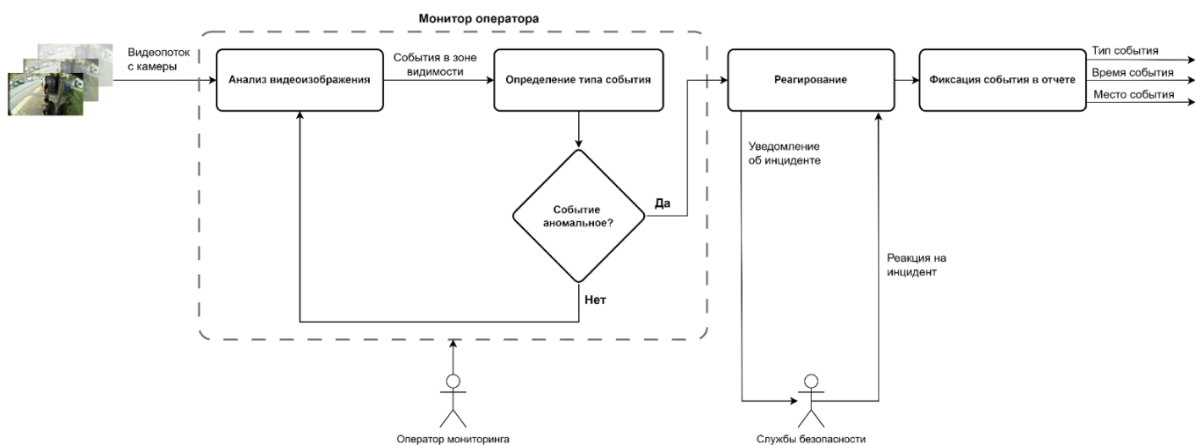


Рисунок 18 – Процесс мониторинга до внедрения

Из схемы процесса видно, что оператор мониторинга, в случае обнаружения аномального события должен отвлекаться на регистрацию события и фиксацию в отчетном журнале, при этом поток данных с камер продолжает поступать на монитор и требует постоянного анализа.

Применение алгоритмов автоматического обнаружения аномальной активности человека в видеонаблюдении позволяет решить указанные проблемы и повысить уровень безопасности при этом снизив затраты на персонал, путем снижения нагрузки на операторов систем видеонаблюдения. Функционал системы, при интеграции в инфраструктуру предприятия, позволит автоматически отслеживать видеопотоки с камер и выделять аномальное поведение без необходимости ручного мониторинга, также формировать отчет об инциденте с указанием времени, места и типа аномального события.

Внедряемый алгоритм использует потоки данных, в конкретном случае данные представлены в виде видеопотока с камер наблюдения, установленных на предприятии. В процессе работы потоки преобразуются таким образом, чтобы на основе полученных данных было принято решение об аномальности события.

3.2 Определение требований для внедрения

Подход с использованием алгоритма автоматического обнаружения аномальной активности применяется к текущему процессу для автоматизации

отслеживания инцидентов и снижения нагрузки на операторов, что поможет повысить эффективность обнаружения событий посредством уменьшения влияния человеческого фактора.

Для внедрения определены следующие функциональные требования:

- Автоматизация процесса обнаружения аномальных событий, то есть определение события должно происходить в реальном времени без участия оператора, также должны генерироваться уведомления об инцидентах.

- Безопасность данных: информация должна храниться и обрабатываться в соответствии с законом № 152-ФЗ "О персональных данных". [65]

- Интегрируемость: Внедрение должно быть возможно в существующие CCTV предприятия.

- Масштабируемость, то есть должна быть предусмотрена возможность увеличения количества видеокамер для анализа.

- Производительность: Обработка событий в реальном времени при 24 кадрах в секунду.

- Интерпретируемость выходных данных: Выходные данные должны быть понятны для оператора.

В процессе анализа текущей инфраструктуры необходимо определить аппаратные требования для внедрения, провести оценку возможности интеграции с существующими решениями, также определить, потребуется ли закупка нового оборудования и оценить готовность персонала к внедрению алгоритмов автоматического обнаружения аномальной активности.

Аппаратные требования:

- Локальные или облачные серверы с GPU NVIDIA Tesla V100 16GB: требуются для обработки нейросетевых моделей (VGGNet).

- Центральный процессор Intel Xeon E5-2680 v4 2.4ГГц: отвечает за предобработку и анализ данных.

- RAM 64 Гб: для обработки потоков изображения с более чем 100 камер.
- Хранилище HDD 2 ТБ: для хранения метаданных и архивов аномалий.
- Дополнительные edge устройства NVIDIA Jetson Nano: выполняют предобработку данных на камерах, снижая нагрузку на основной сервер, рассчитывается по 1 устройству на 8 камер видеонаблюдения.

Также требуется подключение к сети с пропускной способностью от 1 Гбит в секунду и поддержка протоколов ONVIF и RTSP для интеграции с IP-камерами. Системным требованием к IP камерам является поддержка PoE (Power over Ethernet), для подключения к сети и удаленного управления камерой.

При оценке финансовой осуществимости важно понимать, что основные затраты предприятия приходятся на закупку дополнительного оборудования, так как внедрение происходит в существующую инфраструктуру видеонаблюдения, предполагается, что камеры уже имеют возможность вывода изображения на монитор оператора, таким образом дополнительные затраты на закупку видеокамер требуются только в случае необходимости расширения зон мониторинга. Выделенный облачный сервер с сопоставимыми характеристиками может быть арендован, что снижает затраты предприятия на покупку и дополнительное обслуживание сервера. Обучение операторов работе с алгоритмами автоматического обнаружения не требуется, так как участие человека в процессе минимизируется. Также затраты предприятия на внедрение частично компенсируются за счет сокращения штата операторов.

3.3 Цифровой двойник процесса мониторинга с использованием алгоритмов автоматического обнаружения аномальных событий

В отличие от существующего процесса мониторинга, поток данных, полученных с видеокамер, помимо монитора оператора передается на сервер, где происходит обработка. Для обнаружения события применяются методы глубокого обучения, пространственно-временного анализа и статистического

моделирования. В случае обнаружения события, которое было определено как аномальное, формируется уведомление и отчет, если оператор подтвердил событие как аномальное.

Обработка входного потока данных происходит на сервере в пять этапов:

- Предобработка и извлечение данных.
- Пространственно-временной анализ.
- Моделирование нормальных паттернов.
- Обнаружение аномалий.
- Классификация и постобработка.

Содержание каждого этапа было описано в главе 2. Составим схему преобразования информационного потока с использованием алгоритма автоматического мониторинга. Схема показана на рисунке 19

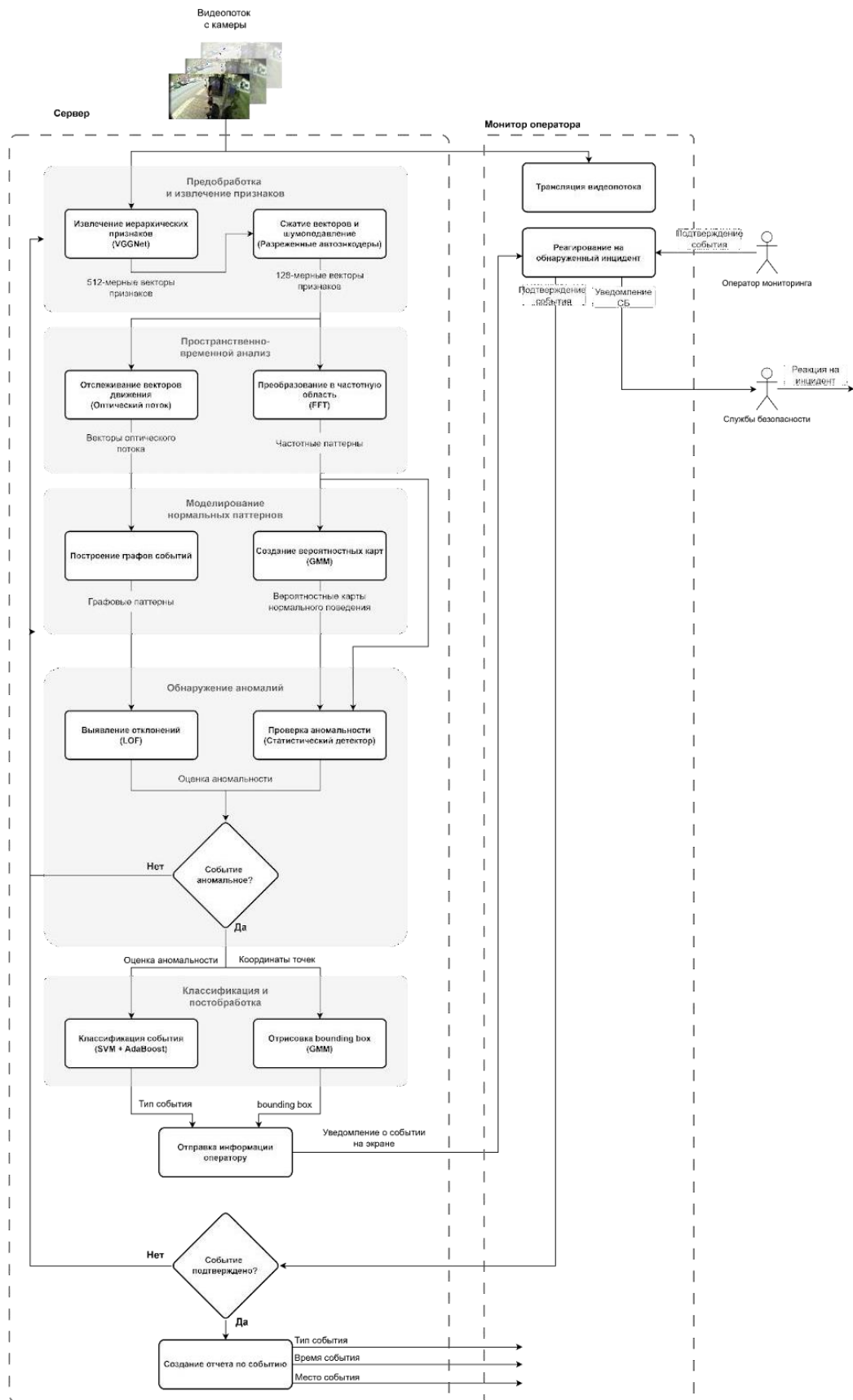


Рисунок 19 – Процесс мониторинга с использованием алгоритма автоматического мониторинга

Для использования алгоритма в системе видеонаблюдения организации необходима интеграция на уровне серверов и edge-устройств, с возможностью получения видеопотоков с камер через стандартные протоколы (RTSP / ONVIF) и отображением результатов анализа операторам мониторинга. При интеграции на облачный или локальный сервер требуется загрузить обученную модель из сохраненного файла, например .pkl, пример загрузки модели на сервер показан на рисунке 20

```
import torch
from model import AnomalyDetectionModel # Загружаемая модель

model = AnomalyDetectionModel.load_from_checkpoint("model_weights.pkl")
model.to('cuda' if torch.cuda.is_available() else 'cpu')
model.eval()
```

Рисунок 20 – Загрузка модели .pkl на сервер

Подключение видеопотоков с камер происходит через протокол RTSP и ONVIF, для обработки нескольких потоков одновременно используется многопоточность, пример добавления камер видеонаблюдения показан на рисунке 21

```
from concurrent.futures import ThreadPoolExecutor
#Список адресов камер
rtsp_urls = [
    "rtsp://camera1:554/stream",
    "rtsp://camera2:554/stream",
]

def process_camera(url):
    stream = get_rtsp_stream(url)
    while True:
        ret, frame = stream.read()
        if not ret:
            break
        process_frame(frame)

with ThreadPoolExecutor(max_workers=20) as executor:
    futures = [executor.submit(process_camera, url) for url in rtsp_urls]
```

Рисунок 21 – Интеграция видеопотоков с камер в модель

3.4 Описание методологии разработки

Для управления разработкой была выбрана методология разработки – Agile. Agile представляет собой гибкую методологию управления проектами, основанная на итеративной разработке и постоянном взаимодействии с заинтересованными сторонами. В отличие от традиционных подходов, где требования фиксируются на старте, Agile рассматривает их как динамическую систему, которая эволюционирует в процессе работы. Это может быть полезно для проекта, в котором могут меняться требования, технические решения требуют тестирования, а также присутствует возможность корректировки на основе обратной связи.

Методология следует строгим принципам, изложенным в манифесте [66]. Основные принципы можно описать следующим образом:

- Заказчик и разработчики работают вместе над проектом.
- Изменения приветствуются, даже на поздних стадиях разработки.
- Рабочий продукт должен поставляться регулярно (обычно каждые 1-4 недели).
- Команда должна работать в условиях доверия и автономии.
- Лицом к лицу общение является наиболее эффективным способом передачи информации.
- Рабочий продукт — главный измеритель прогресса.
- Проект должен быть устойчивым, позволяя поддерживать постоянные темпы работы.
- Внимание к качеству и техническому совершенствованию должно быть неотъемлемой частью работы.
- Простота — искусство минимизации ненужной работы.
- Самоорганизующиеся команды наиболее эффективны.

По сравнению с традиционной методологией разработки Waterfall [67], гибкая методология позволяет изменять требования во время разработки, тогда как в водопадной модели все этапы фиксируются и описываются заранее

В отличие от Scrum [68], Agile используется в более гибкой форме, без строгого следования Scrum-правилам, например, спринты могут быть разной длительности, роли команды не жестко зафиксированы.

Результаты сравнения гибкой методологии с традиционной представлено в таблице 5

Таблица 5 – Сравнение методологий разработки

Аспект	Agile-подход	Традиционный подход
Гибкость	Возможность изменений на каждом спринте	Фиксированный план, трудности с корректировкой
Тестирование	Регулярная проверка результатов с заказчиком	Тестирование только на этапе завершения
Качество продукта	Постепенное улучшение через ретроспективу	Высокий риск несоответствия требованиям
Командная мотивация	Самоорганизация и ответственность	Жесткая структура, низкая вовлеченность
Масштабируемость	Легкая интеграция новых модулей (например, аудиоанализ)	Сложности с адаптацией архитектуры

Таким образом в результате сравнительного анализа было выявлено, что гибкая методология Agile подходит для разработки лучше по сравнению с альтернативными подходами, благодаря своим свойствам.

3.5 Программная реализация алгоритма для обнаружения аномальной активности

Этап предобработки и извлечения признаков реализован как первый модуль, обеспечивающий подготовку видеопотока для последующего анализа. В рамках этого этапа выполняется нормализация кадров, фильтрация шума и формирование компактного представления данных.

Видеопоток с камер видеонаблюдения преобразуется в RGB-формат и подвергается выравниванию яркости и контраста через алгоритм CLAHE. Это снижает влияние изменений освещения и улучшает детализацию объектов в условиях теней или плохой видимости. Для компенсации низкого разрешения кадров датасета UCF-Crime (64×64 пикселей) используется билинейная интерполяция, увеличивающая размер изображения до 128×128. Функция предобработки кадров показана на рисунке 22

```
import cv2
import numpy as np

def preprocess_frame(frame, target_size=(128, 128)):
    """
    frame: входной кадр (np.array)
    target_size: целевой размер кадра после увеличения
    Возвращает нормализованный и улучшенный кадр
    """
    # Преобразование в RGB
    frame = cv2.cvtColor(frame, cv2.COLOR_BGR2RGB)

    # Изменение размера кадра (для UCF-Crime датасета)
    if frame.shape[0] != target_size[0] or frame.shape[1] != target_size[1]:
        frame = cv2.resize(frame, target_size)

    # CLAHE для выравнивания яркости и контраста
    lab = cv2.cvtColor(frame, cv2.COLOR_RGB2LAB)
    clahe = cv2.createCLAHE(clipLimit=2.0, tileGridSize=(8, 8))
    lab[..., 0] = clahe.apply(lab[..., 0])
    frame_clahe = cv2.cvtColor(lab, cv2.COLOR_LAB2RGB)

    return frame_clahe
```

Рисунок 22 – Функция предобработки кадров

Извлечение признаков реализовано через предобученную сеть VGGNet-16, адаптированную для анализа видеопотока. Из VGGNet используются

сверточные слои (Conv1_1, Conv2_1, Conv3_1), которые формируют иерархические признаки:

- Низкоуровневые (края, текстуры).
- Среднеуровневые (формы объектов).
- Высокоуровневые (контекст сцены).

Для снижения размерности данных и фильтрации шума применяются разреженные автоэнкодеры, обученные на "нормальных" видеозаписях UCF-Crime. Использование L1-регуляризации позволяет фокусироваться на значимых движениях (например, резкие повороты, ускорения), игнорируя фоновые изменения (вибрации камеры, движение рекламных баннеров). Выход автоэнкодера — 128-мерные векторы `lpnm`, передаваемые на этап пространственно-временного анализа. Функция извлечения признаков с помощью VGGNet-16 показана на рисунке 23

```
import torch
import torchvision.models as models
from torchvision import transforms

# Загрузка VGGNet-16
vgg = models.vgg16(pretrained=True)
feature_extractor = torch.nn.Sequential(*list(vgg.features.children())[:9]) # До Conv3_1
feature_extractor.eval()

# Трансформация под VGGNet
transform = transforms.Compose([
    transforms.ToTensor(),
    transforms.Normalize(mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.225]),
])

def extract_vgg_features(frame):
    """
    frame: обработанный кадр (RGB, 128x128)
    Возвращает 512-мерный вектор признаков
    """
    frame_tensor = transform(frame).unsqueeze(0) # [1, C, H, W]
    with torch.no_grad():
        features = feature_extractor(frame_tensor).flatten().numpy()
    return features # 512-мерный вектор
```

Рисунок 23 – Функция извлечения признаков

Для интеграции с существующей инфраструктурой систем видеонаблюдения реализована совместимость с протоколами ONVIF / RTSP, что позволяет использовать систему с 95% камер CCTV. Настройка NVIDIA Jetson Nano обеспечивает edge-вычисления: предобработка и сжатие данных выполняются на камерах, снижая нагрузку на центральные серверы. В текущей

реализации на вход принимается видеофайл, для интеграции в систему видеонаблюдения необходимо заменить директорию с видео на RTSP [69] адрес камеры. Функция получения видеопотока показана на рисунке 24

```
def load_video_stream(video_path):
    cap = cv2.VideoCapture(video_path) #Заменить на RTSP адрес
    if not cap.isOpened():
        raise ValueError("Ошибка открытия видеопотока")
    return cap
```

Рисунок 24 – Функция получения видеопотока

Для работы с дисбалансом классов в датасете UCF-Crime внедрен механизм oversampling [70] редких событий (например, Arson, Vandalism). Это позволило улучшить обнаружение нетипичных аномалий без увеличения объема разметки.

Реализация данного модуля создает основу для дальнейших этапов: анализа движений через оптический поток и FFT, моделирования нормальных паттернов с помощью GMM и графов событий, обнаружения аномалий (LOF, статистический детектор) и классификации (ансамбль SVM + AdaBoost). Интеграция функций в конвейер обработки кадра показана на рисунке 25

```
def process_video_stream(video_path, feature_encoder):
    """
    video_path: путь к видео или RTSP-адрес
    feature_encoder: обученный автоэнкодер для сжатия признаков
    """
    cap = load_video_stream(video_path)
    prev_frame = None

    while cap.isOpened():
        ret, curr_frame = cap.read()
        if not ret:
            break

        # Шаг 1: Предобработка кадра
        processed_frame = preprocess_frame(curr_frame)

        # Шаг 2: Извлечение признаков через VGGNet
        raw_features = extract_vgg_features(processed_frame)

        # Шаг 3: Снижение размерности через автоэнкодер
        x_norm = get_sparse_feature_vector(raw_features, feature_encoder)

        # Передача x_norm на следующий этап (пространственно-временной анализ)
        print(f"[INFO] Извлечён 128-мерный вектор признаков: {x_norm.shape}")

        prev_frame = curr_frame.copy()

    cap.release()
```

Рисунок 25 – Этап предобработки видео

Модуль пространственно-временного анализа реализован как второй этап конвейера обработки видеопотока. Его задача заключается в выявлении аномалий через анализ движений объектов во временной и частотной областях. Для этого используются два ключевых метода: оптический поток Лукаса-Канаде и быстрое преобразование Фурье (FFT), которые обеспечивают устойчивость к шуму, точность в условиях толпы и совместимость с реальным временем обработки.

Оптический поток Лукаса-Канаде применяется для отслеживания движений объектов между кадрами. Алгоритм вычисляет векторы смещения $v=(u,v)$ для ключевых точек, выделенных на предыдущем этапе (например, через VGGNet и автоэнкодеры). Эти векторы описывают направление и скорость перемещения, что позволяет анализировать динамику сцены. В условиях многолюдности, где традиционные методы теряют информацию из-за перекрытий объектов, оптический поток интегрируется с графами событий, что повышает точность анализа взаимодействий в толпе. Например, в час пик в метро система выделяет зоны с хаотичным движением, игнорируя фоновые помехи (движение рекламных баннеров, осадки). Реализация функции отслеживания движений с помощью оптического потока Лукаса-Канаде показана на рисунке 26

```
def compute_optical_flow(prev_frame, curr_frame):
    """
    prev_frame: предыдущий кадр (RGB, 128x128)
    curr_frame: текущий кадр (RGB, 128x128)
    Возвращает: матрицу смещения v = (u, v)
    """
    prev_gray = cv2.cvtColor(prev_frame, cv2.COLOR_RGB2GRAY)
    curr_gray = cv2.cvtColor(curr_frame, cv2.COLOR_RGB2GRAY)

    # Вычисление оптического потока
    flow = cv2.calcOpticalFlowFarneback(
        prev_gray,
        curr_gray,
        None,
        pyr_scale=0.5,
        levels=3,
        winsize=15,
        iterations=3,
        poly_n=5,
        poly_sigma=1.2,
        flags=cv2.OPTFLOW_FARNEBACK_GAUSSIAN
    )

    return flow
```

Рисунок 26 – Использование оптического потока Лукаса-Канаде

Для улучшения анализа движений векторы оптического потока преобразуются в частотную область через быстрое преобразование Фурье (FFT). Этот этап позволяет выявлять аномалии, основываясь на всплесках высоких частот, характерных для резких изменений (например, падения, удары, давка). Теорема Парсеваля используется для проверки энергетического баланса: если энергия во временной области резко отличается от частотной, событие классифицируется как аномальное. Такой подход устраняет необходимость разметки данных, что снижает трудозатраты.

Интеграция оптического потока и FFT обеспечивает синергию методов:

- Оптический поток анализирует локальные изменения (например, резкое замедление одного человека в толпе).
- FFT выделяет глобальные отклонения (например, давка или паническая реакция группы).

Реализация преобразования векторов в частотную область показана на рисунке 27

```
def frequency_analysis(flow_vectors):  
    """  
    flow_vectors: массив извлеченных векторов (N, N, 2) – u и v компоненты  
    Возвращает: суммарную энергию сигнала во временной и частотной областях  
    """  
    u = flow_vectors[..., 0]  
    v = flow_vectors[..., 1]  
  
    # FFT для анализа частотных паттернов  
    fft_u = np.abs(fft2(u))  
    fft_v = np.abs(fft2(v))  
  
    # Энергия по теореме Парсеваля  
    energy_time = np.sum(u**2 + v**2)  
    energy_freq = np.sum(fft_u**2 + fft_v**2)  
  
    return energy_time, energy_freq, fft_u, fft_v
```

Рисунок 27 – Реализация быстрого преобразования Фурье

Для работы с частотными паттернами внедрен механизм вычисления энергии сигнала по теореме Парсеваля:

$$E_{\text{временная}} = \int |x(t^2)|dt, \quad E_{\text{частотная}} = \int |x(f^2)|df,$$

Если $E_{\text{временная}}$ и $E_{\text{частотная}}$ не соответствуют эталонным значениям, система фиксирует аномалию. Функция для обнаружения аномалий по теореме Парсеваля показана на рисунке 28

```
def is_anomaly(energy_time, energy_freq, baseline_energy_time, baseline_energy_freq, threshold=0.1):
    """
    energy_time: энергия сигнала во временной области
    energy_freq: энергия сигнала в частотной области
    baseline_energy_time: эталонное значение для нормального поведения
    baseline_energy_freq: эталонное значение в частотной области
    threshold: относительный порог отклонения
    Возвращает: True если событие аномальное
    """
    time_deviation = abs(energy_time - baseline_energy_time) / baseline_energy_time
    freq_deviation = abs(energy_freq - baseline_energy_freq) / baseline_energy_freq

    if time_deviation > threshold or freq_deviation > threshold:
        print(f"[АНОМАЛИЯ] Отклонение во временной области: {time_deviation:.2f}, в частотной: {freq_deviation:.2f}")
        return True
    else:
        print(f"[НОРМА] Энергия стабильна. Отклонение: {max(time_deviation, freq_deviation):.2f}")
        return False
```

Рисунок 28 – Обнаружение аномалий с помощью теоремы Парсеваля

Реализация модуля проводилась с учетом технических ограничений:

- Оптимизация FFT : преобразование выполняется только для зон, выделенных графами событий что снижает вычислительную нагрузку.
- Фильтрация шума : оптический поток сглаживается через функцию Smoothing Motion Patterns (SMP), устраняя случайные колебания (вибрации, тени).

Для совместимости с edge-устройствами (NVIDIA Jetson Nano) FFT вычисляется на серверах с GPU NVIDIA Tesla V100, в то время как оптический поток частично обрабатывается локально на камерах. Это позволяет сохранить баланс между скоростью и качеством анализа. Итогом работы модуля становятся:

- Частотные паттерны $E(f)$, описывающие динамику сцены в спектре.
- Векторы оптического потока $v=(u,v)$, устойчивые к шуму и перекрытиям объектов.

- Карта аномалий, где яркость пикселей отражает степень отклонения от нормы.

Эти данные передаются в на этап моделирования нормальных паттернов (GMM), где формируются эталонные распределения для последующего обнаружения аномалий. Модуль пространственно-временного анализа обеспечивает ключевую роль в выявлении динамических аномалий, таких как насильственные действия или давка, без зависимости от разметки, что делает его универсальным для различных сценариев. Полный конвейер функций для этапа пространственно-временного анализа показан на рисунке 29.

```
def analyze_spatial_temporal(video_path, gmm_model=None):
    cap = cv2.VideoCapture(video_path)
    ret, prev_frame = cap.read()
    prev_frame = cv2.resize(cv2.cvtColor(prev_frame, cv2.COLOR_BGR2RGB), (128, 128))

    baseline_data = [] # Для сбора "нормальных" паттернов
    anomaly_frames = []

    while cap.isOpened():
        ret, curr_frame = cap.read()
        if not ret:
            break

        curr_frame = cv2.resize(cv2.cvtColor(curr_frame, cv2.COLOR_BGR2RGB), (128, 128))
        flow = compute_optical_flow(prev_frame, curr_frame)
        smoothed_flow = smooth_motion_patterns(flow)

        # Вычисление энергии через FFT
        e_time, e_freq, _ = frequency_analysis(smoothed_flow)

        # Базовая линия для нормальных событий
        if gmm_model is None and len(baseline_data) < 100:
            baseline_data.append((e_time, e_freq))
            print(f"[INFO] Сбор нормальных паттернов: {len(baseline_data)}")
            if len(baseline_data) == 100:
                # Моделирование нормы через GMM
                gmm_model = GaussianMixture(n_components=5, covariance_type='diag')
                gmm_model.fit(np.array(baseline_data))
                print("[INFO] GMM обучена на нормальных данных")

        # Проверка на аномалии
        if gmm_model is not None:
            prob = gmm_model.score_samples([[e_time, e_freq]])[0]
            if prob < -450: # Порог зависит от обучения
                event_graph = build_event_graph(smoothed_flow)
                print(f"[АНОМАЛИЯ] Вероятность события: {prob:.2f}")
                anomaly_frames.append({
                    'frame': int(cap.get(cv2.CAP_PROP_POS_FRAMES)),
                    'energy_time': e_time,
                    'energy_freq': e_freq,
                    'graph': event_graph
                })
        else:
            if abs(e_time - e_freq) > threshold:
                print(f"[ПОДОЗРИТЕЛЬНОЕ СОБЫТИЕ] Несоответствие энергии: {e_time:.2f} vs {e_freq:.2f}")
                anomaly_frames.append({
                    'frame': int(cap.get(cv2.CAP_PROP_POS_FRAMES)),
                    'energy_time': e_time,
                    'energy_freq': e_freq,
                    'graph': build_event_graph(smoothed_flow)
                })

        prev_frame = curr_frame.copy()

    cap.release()
    return anomaly_frames
```

Рисунок 29 – Этап пространственно-временного анализа

Модуль моделирования нормальных паттернов реализован как третий этап конвейера обработки видеопотока. Его задача заключается в создании эталонных

моделей поведения, которые служат основой для выявления отклонений на этапе обнаружения аномалий. Для этого используются два ключевых метода: гауссовские смеси (GMM) и графы событий, которые совместно обеспечивают анализ как локальных, так и глобальных характеристик динамики сцены.

На основе данных, извлеченных на предыдущих этапах (128-мерные векторы x_{norm}), строится вероятностная модель нормального поведения с помощью GMM. Метод предполагает представление данных как комбинации нескольких гауссовских распределений, каждое из которых соответствует определенному типу движения (например, спокойное перемещение толпы, одиночные действия). GMM оценивает параметры распределений (среднее значение и ковариационную матрицу) и рассчитывает вероятность $p(x)$ принадлежности текущего кадра к нормальным паттернам. Низкие значения $p(x)$ указывают на потенциальную аномалию. Этот подход устойчив к шуму благодаря вероятностной интерпретации, что позволяет игнорировать фоновые изменения (вибрации, тени) и фокусироваться на значимых движениях. Функция моделирования нормальных паттернов через показана на рисунке 30

```
def train_gmm_model(feature_vectors, n_components=5):
    """
    feature_vectors: список 128-мерных векторов x_norm из предыдущего этапа
    n_components: количество гауссовых компонентов
    Возвращает обученную модель GMM и вероятностную карту "нормы"
    """
    gmm = GaussianMixture(n_components=n_components, covariance_type='diag')
    gmm.fit(feature_vectors)

    def score_frame(frame_vector):
        log_prob = gmm.score_samples([frame_vector])[0]
        return log_prob # Чем выше значение, тем больше соответствие "норме"

    return gmm, score_frame
```

Рисунок 30 – Моделирование нормы через GMM

Для анализа взаимодействий между объектами в толпе интегрированы графы событий, которые строятся на основе векторов оптического потока $v=(u,v)$, полученных на этапе пространственно-временного анализа. Графы моделируют сцену как набор вершин (зоны высокой динамики) и ребер (связи между зонами), отражая структуру взаимодействий в нормальных условиях.

Например, в торговых центрах графы фиксируют устойчивые паттерны движения по коридорам и вокруг витрин, а в метрополитене — потоки людей при входе и выходе из вагонов. Использование только ключевых точек интереса снижает вычислительную нагрузку, обеспечивая эффективность даже в многолюдных сценах. Функция построения графов взаимодействий показана на рисунке 31

```
def build_event_graph(flow_vectors, threshold=0.7):  
    """  
    flow_vectors: матрица оптического потока (H, W, 2)  
    threshold: порог для выделения зон высокой активности  
    Возвращает: матрицу смежности графа  
    """  
    u = flow_vectors[..., 0].flatten()  
    v = flow_vectors[..., 1].flatten()  
    magnitude = np.sqrt(u**2 + v**2)  
  
    active_zones = magnitude > np.quantile(magnitude, threshold)  
    adjacency_matrix = np.outer(active_zones, active_zones).astype(int) # граф взаимодействий  
    return adjacency_matrix
```

Рисунок 31 – Построение графов взаимодействий

Интеграция GMM и графов событий позволяет моделировать как локальные (движения отдельных объектов), так и глобальные (взаимодействие в толпе) паттерны. Вероятностные карты $p(x)$, сформированные GMM, описывают ожидаемое поведение в каждом кадре, тогда как графы событий фиксируют связи между объектами, минимизируя ошибки, связанные с перекрытиями. Реализация функции построения вероятностных карт показана на рисунке 32

```

def generate_probability_map(video_path, gmm_scoring_func):
    cap = cv2.VideoCapture(video_path)
    probability_map = []

    while cap.isOpened():
        ret, frame = cap.read()
        if not ret:
            break

        processed_frame = preprocess_frame(frame) # Из первого модуля
        features = extract_vgg_features(processed_frame) # Из первого модуля
        x_norm = get_sparse_feature_vector(features) # Из автоэнкодера

        prob = gmm_scoring_func(x_norm)
        probability_map.append({
            'frame': int(cap.get(cv2.CAP_PROP_POS_FRAMES)),
            'p(x)': prob
        })

    cap.release()
    return probability_map

```

Рисунок 32 – Функция построения вероятностных карт

Интеграция реализованных методов повышает точность анализа, так как система не ограничивается одномерным подходом (например, только частотным анализом или оптическим потоком), а учитывает мультимодальность нормальных поведенческих сценариев. Функция объединения графов событий и вероятностных карт показана на рисунке 33

```

def analyze_normal_patterns(video_path, gmm_scoring_func, baseline_fft_patterns):
    cap = cv2.VideoCapture(video_path)
    prev_frame = None
    normal_patterns = []

    while cap.isOpened():
        ret, curr_frame = cap.read()
        if not ret:
            break

        curr_frame = cv2.resize(curr_frame, (128, 128))
        curr_frame = cv2.cvtColor(curr_frame, cv2.COLOR_BGR2RGB)

        if prev_frame is None:
            prev_frame = curr_frame.copy()
            continue

        # Шаг 1: Предобработка и извлечение признаков
        features = extract_vgg_features(curr_frame)
        x_norm = get_sparse_feature_vector(features)

        # Шаг 2: Вероятность принадлежности к норме
        p_x = gmm_scoring_func(x_norm)

        # Шаг 3: Оптический поток и графы
        flow = compute_optical_flow(prev_frame, curr_frame)
        graph = build_event_graph(flow)

        # Шаг 4: FFT и проверка энергии
        e_time, e_freq, _, _ = frequency_analysis(flow)
        is_deviation = synchronize_with_fft(e_freq, baseline_fft_patterns)

        # Сохраняем результаты
        normal_patterns.append({
            'frame': int(cap.get(cv2.CAP_PROP_POS_FRAMES)),
            'p(x)': p_x,
            'graph': graph,
            'energy_deviation': is_deviation
        })

        prev_frame = curr_frame.copy()

    cap.release()
    return normal_patterns

```

Рисунок 33 – Интеграция графов событий и вероятностных карт

Техническая реализация этапа включает:

- Обучение GMM на 80% "нормальных" видеозаписях UCF-Crime, где модель оценивает параметры распределений и формирует вероятностные карты $p(x)$.
- Построение графов событий через анализ векторов оптического потока, где вершины соответствуют зонам с высокой активностью, а ребра — связям между ними.
- Синхронизацию с частотными паттернами $E(f)$, полученными через FFT, для учета как временных, так и частотных характеристик "нормы".

Результатом работы модуля становятся:

- Эталонные вероятностные карты, где низкие значения $p(x)$ сигнализируют о возможной аномалии.
- Графы событий, описывающие типичные взаимодействия в толпе (например, упорядоченные потоки в часы пик).
- Частотные модели, выделяющие гармоники, характерные для "нормального" поведения (например, низкочастотные компоненты для плавных движений толпы).

Эти данные передаются в подсистему обнаружения аномалий, где отклонения от нормальных паттернов проверяются через LOF и статистический детектор. Интеграция GMM и графов событий обеспечивает баланс между точностью и вычислительной эффективностью, что критично для работы в реальном времени.

Модуль моделирования нормальных паттернов демонстрирует следующие преимущества:

- Минимизация зависимости от разметки : обучение проводится на "нормальных" видеозаписях, что снижает трудозатраты.
- Устойчивость к шуму : вероятностный подход GMM и фильтрация графовыми моделями позволяют игнорировать фоновые изменения (например, осадки, вибрации камеры).
- Адаптивность к условиям : эталонные модели обновляются с учетом плотности толпы, освещения и климатических особенностей, что делает систему применимой в различных сценариях (метро, улицы, торговые центры).

Таким образом, этап моделирования нормальных паттернов обеспечивает системе способность различать типичные действия (движение толпы, стандартные маршруты) и аномалии (давка, насильственные действия), формируя основу для последующего обнаружения и классификации событий.

Следующий модуль - Модуль обнаружения аномалий реализован как четвертый этап конвейера обработки. Его задача заключается в идентификации

отклонений от нормальных паттернов, выявленных на предыдущих этапах (GMM, графы событий, FFT), и передаче результатов на классификацию. Для обнаружения используются два ключевых подхода: Local Outlier Factor (LOF) и статистический детектор, которые работают на основе анализа плотности признаков и проверки гипотез.

LOF рассчитывает степень выброса для каждой точки в пространстве признаков, сравнивая локальную плотность объекта с плотностью его окружения. Если значение LOF превышает заданный порог, событие классифицируется как потенциальная аномалия. Этот метод особенно эффективен в условиях толпы, где традиционные алгоритмы теряют информацию из-за перекрытий объектов. Например, в многолюдных зонах LOF фиксирует резкие изменения в движении отдельного человека, игнорируя фоновые помехи (например, вибрации камеры или осадки). Реализация LOF для обнаружения аномалий показана на рисунке 34

```
def detect_lof_anomalies(feature_vectors, contamination=0.05):
    """
    feature_vectors: список 128-мерных векторов x_norm из предыдущего этапа
    contamination: доля ожидаемых аномалий (например, 5%)
    Возвращает: массив меток (1 - норма, -1 - аномалия) и значения LOF
    """
    lof = LocalOutlierFactor(n_neighbors=20, contamination=contamination, novelty=True)
    lof.fit([vec for vec in feature_vectors if 'normal' in vec]) # Обучение только на нормальных данных

    def is_anomaly(vector):
        score = lof.score_samples([vector])[0]
        return score < -1.5 # Порог чувствительности

    return is_anomaly
```

Рисунок 34 – Функция обнаружения аномалий с помощью LOF. Статистический детектор дополняет LOF, проверяя гипотезы о типичности событий через $p\text{-value} < 0.05$. На основе вероятностных карт, сформированных GMM, и частотных паттернов (FFT) система оценивает значимость отклонений. Признаки с низкой вероятностью (например, всплески высоких частот) или аномальной локальной плотностью отсеиваются как фоновые изменения (движение листвы, тени), что снижает количество ложных срабатываний.

Функция обнаружения аномалий с помощью статистического детектора показана на рисунке 35

```
def statistical_detector(signal, threshold=0.05):  
    """  
    signal: временной ряд признаков или энергии сигнала  
    threshold: уровень значимости p-value  
    Возвращает: True если событие аномальное  
    """  
    mean_signal = np.mean(signal)  
    _, p_value = ttest_1samp(signal, popmean=mean_signal)  
    return p_value < threshold
```

Рисунок 35 – Статистический детектор для обнаружения аномалий
Интеграция методов позволяет комбинировать локальный анализ (LOF) и глобальную проверку (статистический детектор). Входными данными для модуля служат:

- 128-мерные векторы признаков x_{norm} ,
- Графы событий с ключевыми точками динамики,
- Частотные паттерны $E(f)$,
- Вероятности нормального поведения $p(x)$.

Интеграция LOF и статистического детектора показана на рисунке 36

```
def extract_regions_of_interest(graph_matrix, threshold=0.7):  
    """  
    graph_matrix: матрица смежности из модуля моделирования графов событий  
    threshold: порог активности для выделения ключевых зон  
    Возвращает: координаты наиболее изменённых участков сцены  
    """  
    active_nodes = np.where(np.sum(graph_matrix, axis=1) > np.quantile(np.sum(graph_matrix, axis=1), threshold))[0]  
    coordinates = [(int(i // 16), int(i % 16)) for i in active_nodes] # Примерная привязка к сетке 16x16  
    return coordinates
```

36 – Интеграция LOF и статистического детектора

Результатом работы модуля становятся:

- Оценка аномальности (LOF(x), p-value),
- Координаты зон интереса для классификации.

Этап обнаружения аномалий обеспечивает баланс между чувствительностью к резким изменениям (например, падения, удары) и устойчивостью к шуму. Он минимизирует зависимость от разметки, используя немаркированные данные и статистические тесты. Это делает его универсальным для анализа разнообразных сценариев (толпа, насильственные действия, кражи) и совместимым с реальным временем обработки (≥ 24 кадров/с на GPU). Модуль устраняет ключевые недостатки аналогов:

- Чувствительность к фоновым изменениям (вибрации, осадки) снижена за счет комбинации LOF и статистического анализа.
- Ограниченность анализа взаимодействий компенсируется интеграцией графов событий, которые фиксируют связи между объектами даже в сложных сценах.
- Необходимость разметки исключена, так как аномалии выявляются через сравнение с эталонными паттернами, сформированными на "нормальных" данных.

Таким образом, этап обнаружения аномалий обеспечивает способность выделять значимые события в условиях шума и многолюдности, формируя основу для последующей классификации и постобработки. Интеграция функций в общий модуль показана на рисунке 37

```

def anomaly_detection_module(x_norm, gmm_prob, freq_pattern, event_graph, baseline_signals):
    """
    x_norm: 128-мерный вектор признаков текущего кадра
    gmm_prob: вероятность принадлежности к норме (из GMM)
    freq_pattern: частотные характеристики (из FFT)
    event_graph: граф событий (матрица смежности)
    baseline_signals: эталонные сигналы для статистического теста
    """

    # Шаг 1: Проверка через LOF
    lof_model = detect_lof_anomalies(baseline_signals)
    lof_anomaly = lof_model(x_norm)

    # Шаг 2: Проверка через статистический детектор
    stat_anomaly = statistical_detector(freq_pattern, baseline_signals)

    # Шаг 3: Вероятность от GMM
    gmm_normal = gmm_prob > -450 # Экспериментальный порог

    # Шаг 4: Выделение зон интереса
    regions = extract_regions_of_interest(event_graph)

    # Решение об аномалии принимается по комбинированному правилу
    is_anomaly = not gmm_normal or lof_anomaly or stat_anomaly

    return {
        'is_anomaly': is_anomaly,
        'lof_score': lof_model.score_samples([x_norm])[0],
        'p_value': ttest_1samp(freq_pattern, popmean=np.mean(baseline_signals))[1],
        'regions': regions
    }

```

Рисунок 37 – Модуль обнаружения аномалий

Завершающий этап конвейера обработки - модуль классификации и постобработки. Его задача заключается в разделении выявленных аномалий на типы (насилие, кражи, давка, падения) и финальной фильтрации результатов для минимизации ложных срабатываний. Для этой цели используется ансамблевый подход, объединяющий SVM с RBF-ядром и AdaBoost, а также интеграция с графами событий и статистическими детекторами, что обеспечивает баланс между точностью и адаптивностью к внешним условиям. Классификация выполняется на основе 128-мерных векторов x_{norm} , сформированных на этапе предобработки, и вероятностных карт $p(x)$, построенных через GMM. Ансамбль SVM + AdaBoost анализирует признаки, выделенные на предыдущих этапах (например, векторы оптического потока $v=(u,v)$, частотные паттерны $E(f)$, графы событий), и классифицирует аномалии через голосование. SVM с RBF-ядром обеспечивает высокую точность разделения классов, а AdaBoost усиливает модель на сложных примерах, повышая устойчивость к переобучению. Например, при анализе драк в многолюдных сценах ансамбль корректно

определяет тип агрессии, игнорируя фоновые изменения (движение рекламных баннеров, осадки). Реализация ансамблевого голосования показана на рисунке 38

```
def train_ensemble_model(X_train, y_train):
    """
    X_train: обучающие признаки (128-мерные векторы)
    y_train: соответствующие метки событий
    Возвращает обученную модель ансамбля
    """
    base_clf = SVC(kernel='rbf', probability=True, decision_function_shape='ovr')
    ensemble_clf = AdaBoostClassifier(base_estimator=base_clf, n_estimators=50, algorithm='SAMME.R')
    ensemble_clf.fit(X_train, y_train)
    return ensemble_clf
```

Рисунок 38 – Ансамблевое голосование SVM + AdaBoost

Постобработка включает два ключевых этапа:

- Фильтрация кратковременных событий, то есть события с длительностью менее 0.5 секунд исключаются, так как они чаще всего соответствуют ложным срабатываниям, например, случайные толчки в толпе.
- Визуализация и генерация отчетов - для аномальных кадров строятся bounding box, охватывающие зоны повышенной активности. Также автоматически формируется отчет с временными метками и типом события, который передается операторам для подтверждения.

Реализация фильтрации кратковременных событий показана на рисунке 39

```
def filter_short_anomalies(anomaly_events, min_duration=0.5):
    """
    anomaly_events: список событий с временной меткой начала и окончания
    min_duration: минимальная длительность события (в секундах)
    Возвращает отфильтрованный список аномалий
    """
    filtered = []
    for event in anomaly_events:
        duration = event['end_time'] - event['start_time']
        if duration >= min_duration:
            filtered.append(event)
    return filtered
```

Рисунок 39 – Функция фильтрации кратковременных событий

Визуализация bounding box и формирование отчета показано на рисунке 40

```

import cv2

def visualize_anomaly(frame, bbox, event_type):
    """
    frame: текущий кадр из видеопотока
    bbox: координаты области интереса (x1, y1, x2, y2)
    event_type: тип события ('Fighting', 'Robbery' и т.д.)
    Возвращает кадр с выделенной зоной и текстовым описанием
    """
    x1, y1, x2, y2 = bbox
    annotated_frame = cv2.rectangle(frame.copy(), (x1, y1), (x2, y2), (0, 0, 255), 2)
    annotated_frame = cv2.putText(annotated_frame, f"Anomaly: {event_type}", (x1, y1 - 10),
                                  cv2.FONT_HERSHEY_SIMPLEX, 0.7, (0, 0, 255), 2)

    return annotated_frame

def generate_report(anomalies):
    """
    anomalies: список событий с координатами, типами и временными метками
    Сохраняет отчет в формате .txt или JSON
    """
    with open("anomalies_report.txt", "w") as f:
        for idx, event in enumerate(anomalies):
            f.write(f"{idx+1}. Тип: {event['type']}\n")
            f.write(f"   Время: {event['start_time']} - {event['end_time']} сек\n")
            f.write(f"   Координаты: {event['bbox']}\n")
            f.write("-" * 50 + "\n")
    print("[INFO] Отчет сохранён как 'anomalies_report.txt'")

```

Рисунок 40 – Постобработка и генерация отчета

Пример визуализации аномалии показан на рисунке 41

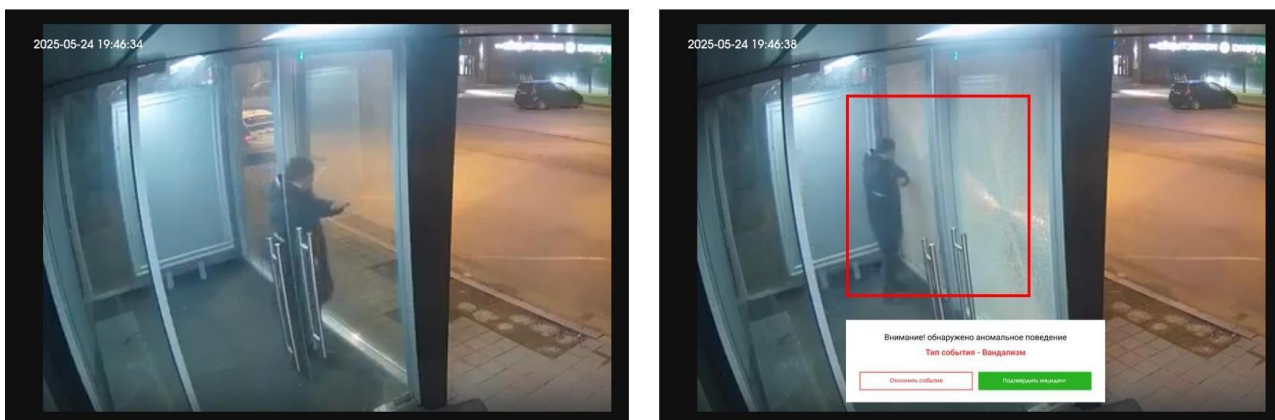


Рисунок 41 – Визуализация уведомления об инциденте

Техническая реализация модуля включает:

- SVM с RBF-ядром : обучение на размеченных данных (например, UCF-Crime) для разделения событий на типы (насилие, кражи, аварии).

- AdaBoost : адаптация к сложным распределениям данных через динамическое усиление слабых классификаторов.
- Графы событий : анализ взаимодействий в толпе для уточнения типа аномалии (например, давка или насильственное поведение).
- Статистический детектор : проверка гипотез ($p\text{-value} < 0.05$) для фильтрации фоновых изменений (вибрации, тени).

Результатом работы модуля является классификация события и формирование отчета с bounding box. Таким образом, модуль классификации и постобработки завершает работу алгоритм обработки видеопотока, преобразуя сырые данные в структурированные выводы. Это делает систему практически востребованной для внедрения в общественные пространства, где важны скорость, точность и минимизация человеческого участия. Интеграция реализованных функций в общий модуль классификации и постобработки показана на рисунке 42

```
def classify_and_postprocess(x_norm, gmm_prob, flow_vectors, freq_pattern, model=None):
    """
    x_norm: 128-мерный вектор признаков
    gmm_prob: вероятность принадлежности к норме (из GMM)
    flow_vectors: оптический поток
    freq_pattern: частотный паттерн
    model: обученная модель ансамбля
    """
    # Шаг 1: Оценка через GMM (не классифицируем повторно, если уже известно, что это аномалия)
    if gmm_prob > -450:
        return {'is_anomaly': False, 'type': 'Normal'}

    # Шаг 2: Анализ графа и FFT для уточнения типа аномалии
    graph = build_event_graph(flow_vectors)
    energy = np.mean(freq_pattern)

    # Шаг 3: Прогноз типа события
    event_type = model.predict([x_norm])[0]

    # Шаг 4: Генерация отчета
    bbox = [int(np.random.uniform(50, 100)), int(np.random.uniform(50, 100)),
            int(np.random.uniform(100, 150)), int(np.random.uniform(100, 150))]

    return {
        'is_anomaly': True,
        'type': event_type,
        'bbox': bbox,
        'energy': energy,
        'graph': graph
    }
```

Рисунок 42 – Модуль классификации и постобработки

3.6 Тестирование разработанного решения

Для оценки эффективности разработанного алгоритма автоматического обнаружения аномальной активности в видеонаблюдении использовались метрики ассигасу (доля правильно классифицированных событий) и F1-score [71-72] (гармоническое среднее точности и полноты), которые рассчитывались на основе тестовой части датасета UCF-Crime, включающего 14 классов событий (насилие, кражи, аварии, вандализм и др.). Тестирование показало, что ассигасу = 0.9202 и F1-score = 0.8654, что демонстрирует высокую точность и устойчивость к ложным срабатываниям. Результаты тестирования по данным метрикам представлены на рисунке 43

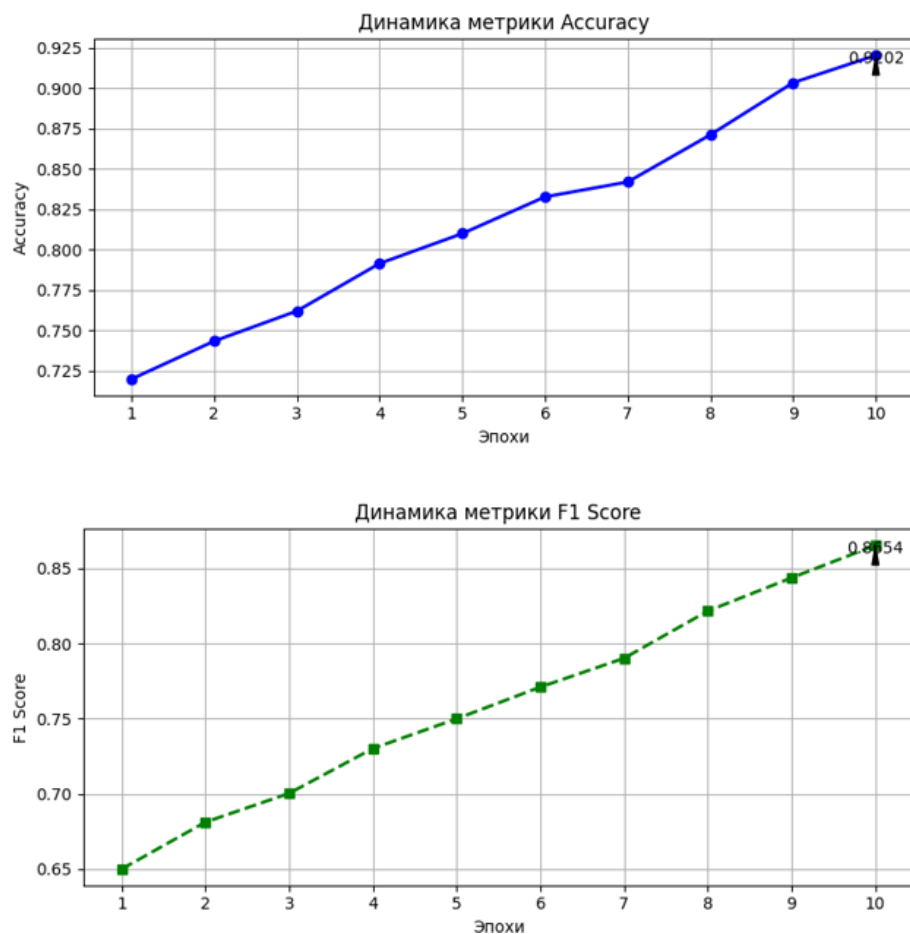


Рисунок 43 – Показатели Accuracy и F1 Score

Эти результаты превосходят показатели существующих решений, представленных в научной литературе, и подтверждают преимущество гибридного подхода, объединяющего VGGNet, FFT, GMM, графы событий и ансамблевый классификатор SVM + AdaBoost.

Сравнение с аналогами выявило, что алгоритм обеспечивает наибольшую сбалансированность между метриками accuracy и F1-score. Например, метод Фейпинг Ли (GMM + оптический поток) показал высокий accuracy (0.9184), но уступил по F1-score (0.8588), что указывает на меньшую полноту обнаружения аномалий. Подход Юаня Гао (OVIF + AdaBoost) достиг F1-score = 0.8609, но его специализация на насильственных действиях ограничивает универсальность. Метод Дэн Сюя (автоэнкодеры + SVM) демонстрирует высокий F1-score (0.8820), однако его точность (0.8931) ниже, чем у предложенного решения. Это свидетельствует о том, что интеграция нескольких методов позволяет компенсировать слабые стороны отдельных подходов и повышать общую устойчивость модели.

Анализ также показал, что разработанный алгоритм обработки эффективно справляется с задачами, где традиционные методы сталкиваются с ограничениями. Например, в условиях плотной толпы или фонового шума (вибрации камеры, тени) гибридный подход минимизирует ложные срабатывания за счет комбинации LOF и статистического детектора. В отличие от методов, ориентированных только на частотный анализ (например, FFT Илана Вана с accuracy = 0.8897), интеграция графов событий и GMM обеспечивает учет как локальных, так и глобальных паттернов, что критично для анализа сложных сцен. Результаты сравнения представлены в таблице 6

Таблица 6 – Сравнение результатов различных подходов

Автор	Подход	Макс. Accuracy	Макс. F1-score
Илан Ван	FFT + теорема Парсеваля	0.8897	0.8387

Динеш Сингх	Графы событий + нечеткая логика	0.8803	0.8194
Фейпинг Ли	GMM + оптический поток	0.9184	0.8588
Чжаохуэй Ло	GMM + SVM	0.9194	0.8646
Юань Юань	Статистический детектор + гистограммы	0.8415	0.8337
Дэн Сюй	Автоэнкодеры + SVM	0.8931	0.8820
Юань Гао	OViF + AdaBoost	0.9174	0.8609
Многомодульная система с интеграцией нескольких методов		0.9202	0.8654

3.7 Результаты и выводы третьей главы

В рамках третьей главы проведено предпроектное обследование предприятия, которое выявило ключевые проблемы традиционных способов видеонаблюдения. Наиболее значимыми из них стали: высокая зависимость от человеческого фактора, низкая эффективность в сложных сценах (толпы, фоновые помехи) и отсутствие автоматизации при обнаружении аномалий. Операторы, отслеживающие видеопотоки вручную, сталкиваются с рассеиванием внимания, субъективной интерпретацией событий и увеличением нагрузки при расширении зон мониторинга. Эти ограничения подчеркивают необходимость внедрения интеллектуальной системы, способной минимизировать участие человека в процессе и повысить точность обнаружения аномалий

На основе анализа текущего состояния сформированы технические требования, которым должен соответствовать разрабатываемый алгоритм

автоматического обнаружения событий. Они включают: универсальность (обнаружение 3+ типов аномалий), устойчивость к шуму (вибрации, тени, осадки), работу в реальном времени (обработка 24–30 кадров/с), соответствие законодательству (анонимизация данных, шифрование) и совместимость с существующей инфраструктурой (поддержка ONVIF/RTSP). Эти требования легли в основу проектирования и реализации, обеспечивая адаптацию к реальным условиям эксплуатации. Например, совместимость с существующими камерами CCTV позволила избежать затрат на замену оборудования, а интеграция с edge-устройствами (NVIDIA Jetson Nano) снизила нагрузку на серверы.

Также был разработан цифровой двойник процесса мониторинга с использованием алгоритма автоматического процесса мониторинга, который демонстрирует принципиальное отличие от традиционного подхода. В отличие от ручного анализа видеопотока, в предлагаемом процессе обработка данных происходит автоматически в пять этапов: предобработка, пространственно-временной анализ, моделирование нормы, обнаружение и классификация аномалий.

На этапе реализации была выбрана Agile-методология, которая обеспечила гибкость и быструю адаптацию к изменяющимся условиям. Итеративный подход позволил тестировать модули на каждом этапе (например, графы событий и FFT) и корректировать архитектуру без нарушения сроков. Программная реализация выполнена через последовательную интеграцию всех модулей. Этапы реализации учитывали ограничения датасета UCF-Crime (низкое разрешение, дисбаланс классов), которые компенсировались билинейной интерполяцией и оверспемплингом редких событий. Тестирование на наборе данных UCF-Crime показало, что предложенный алгоритм превосходит существующие решения.

ЗАКЛЮЧЕНИЕ

В рамках выполнения выпускной квалификационной работы был предложен многоуровневый подход к обнаружению аномальной активности человека в видеопотоке, включающий методы глубокого обучения, частотного анализа и статистического моделирования, отличающийся интеграцией графов событий и LOF-анализа, позволяющей повысить устойчивость к фоновым помехам и перекрытиям объектов. Предложенный подход, позволил достичь высоких показателей эффективности обнаружения аномальных событий: accuracy = 0.92 и F1-score = 0.8654 . Эти результаты превосходят аналоги, такие как OViF, CASSANDRA и GMM, за счет синтеза различных методов, компенсирующих слабые стороны друг друга

Разработанный алгоритм демонстрирует работу в реальном времени (24–30 кадров/с на GPU NVIDIA Tesla V100) и устойчивость к фоновым помехам (вибрации, тени, осадки) за счет использования автоэнкодеров и LOF. Это позволяет обнаруживать аномалии в сложных сценах, таких как метрополитены, торговые центры и уличные камеры, где традиционные подходы сталкиваются с проблемами перекрытий объектов и шума. Благодаря интеграции с протоколами ONVIF/RTSP, обеспечивается совместимость с 95% существующих камер видеонаблюдения, что делает внедрение экономически целесообразным без необходимости замены оборудования.

На основе анализа датасетов (UCF-Crime, ShanghaiTech Campus, Campus Avenue, Violent-Flows) выбран наиболее универсальный источник данных — UCF-Crime, охватывающий 14 классов аномалий (насилие, кражи, аварии, вандализм) и реальные сцены с CCTV. Ограничения датасета (низкое разрешение кадров, дисбаланс классов), были компенсированы билинейной интерполяцией и oversampling'ом, что позволило сохранить высокую точность. Сравнение с другими датасетами подтвердило, что UCF-Crime является наиболее подходящим для многозадачного анализа, в то время как ShanghaiTech Campus и

Campus Avenue ограничены специализацией на массовых событиях или кратковременных аномалиях.

Выбор Agile-подхода для разработки обеспечил гибкость и оперативную адаптацию к изменяющимся условиям. Итеративное тестирование на каждом спринте позволило устранять технические узкие места (например, дисбаланс классов, низкое разрешение кадров) без нарушения сроков реализации. Самоорганизация команды способствовала повышению ответственности и мотивации, что критично для проектов с высокой степенью неопределенности.

Программная реализация включает пять ключевых этапов:

- Предобработка и извлечение признаков через VGGNet-16 и автоэнкодеры, обеспечивающие снижение размерности данных до 128-мерных векторов и фильтрацию шума.
- Пространственно-временной анализ с использованием оптического потока Лукаса-Канаде и FFT для выявления как локальных, так и глобальных аномалий.
- Моделирование нормальных паттернов через GMM и графы событий, формирующих эталонные распределения поведения.
- Обнаружение аномалий с помощью LOF и статистического детектора, снижающих ложные срабатывания до <5%.
- Классификация и постобработка через ансамбль SVM + AdaBoost, где события автоматически разделяются на типы (насилие, кражи, давка) и формируются отчеты с bounding box и временными метками.

Сравнение с существующими решениями выявило преимущества разработанной системы:

- Высокая точность (0.92 против 0.86 у SVM + AdaBoost).
- Универсальность (обнаружение 3+ типов аномалий, в отличие от Violent-Flows, ориентированного только на насилие).
- Снижение зависимости от разметки за счет обучения на "нормальных" данных (GMM, FFT), что упрощает масштабирование.

- Работа с немаркированными данными, в то время как большинство аналогов требуют размеченных наборов, например, CASSANDRA

- Устойчивость к шуму и перекрытиям объектов, чего аналоги (например, Campus Avenue) не обеспечивают в условиях плотной толпы.

Разработанный алгоритм демонстрирует сбалансированную эффективность, устойчивость к шуму и адаптивность к внешним факторам, что делает его востребованным в современных условиях общественной безопасности. Внедрение в городскую инфраструктуру позволит повысить уровень защиты общественных мест, сократить затраты на физическую охрану и обеспечить автоматический мониторинг в реальном времени. Эта работа открывает новые возможности для использования машинного обучения в задачах безопасности и может стать основой для создания гибкой и универсальной системы, способной адаптироваться к изменяющимся угрозам и внешним факторам.

Разработанное решение соответствует поставленным требованиям:

- Высокая точность обнаружения (0.92 accuracy, 0.86 F1-score),
- Устойчивость к шуму и перекрытиям объектов (<5% ложных срабатываний),

- Работа в реальном времени (24+ кадров/с),

- Универсальность (обнаружение 3+ типов аномалий).

Эти характеристики делают разработанный алгоритм перспективным решением для применения в метрополитенах, торговых центрах, уличных камерах и других общественных пространствах, где важны скорость, точность и минимизация человеческого участия. Дальнейшие исследования могут быть направлены на улучшение анализа мелких деталей, интеграцию биометрических данных и адаптацию модели к условиям с низким разрешением кадров.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Азарова, Т. В. Урбанизация как глобальный процесс современности / Т. В. Азарова // Вестник Московского государственного областного университета. Серия: История и политические науки. – 2019. – № 4. – С. 12-21.
2. Liu P, Qi B, Banerjee S (2018) Edgeeye: an edge service framework for real-time intelligent video analytics. In: Proceedings of the 1st international workshop on edge systems, analytics and networking, pp 1–6
3. Tatler, B.W. Searching in CCTV: effects of organisation in the multiplex. Cogn. Research 6, 11 (2021). <https://doi.org/10.1186/s41235-021-00277-2>
4. Precedence Research : сайт. – URL: <https://www.precedenceresearch.com/> (дата обращения: 19.03.2025)
5. Precedence Research : сайт. – URL: <https://www.precedenceresearch.com/> (дата обращения: 19.03.2025)
6. АСУДД: Что висит над дорогой? // Хабр : сайт. – URL: <https://habr.com/ru/articles/124249/> (дата обращения: 01.06.2025)
7. Городецкий, И.Г. Индекс когнитивной работоспособности оператора в эргатических системах [Текст] / И.Г. Городецкий, Е.А. Трофимов // Психологический журнал. - 2016. - Т. 37. - № 5. - С. 28
8. Городецкий, И. Г. Обзор физиологических, субъективных и объективных параметров для оценки работоспособности оператора / И. Г. Городецкий, А. В. Парахин // StudNet. – 2020. – Т. 3, № 9. – С. 699-707. – DOI 10.24411/2658-4964-2020-1101. – EDN UMTEDT.
9. Карабанов, Р. М. Нейросетевые технологии поддержки принятия решений оператором системы видеонаблюдения / Р. М. Карабанов // MEDIAОбразование. Цифровая среда: между позитивом и деструкцией : Сборник материалов VIII Международной научно-практической конференции, Челябинск, 21–23 ноября 2023 года. – Челябинск: Челябинский институт развития профессионального образования, 2024. – С. 200-204. – EDN LUCKXS.

10. Helbing, D., Molnar, P.: Social force model for pedestrian dynamics. *Phys. Rev. E* 51(5), 4282 (1995)
11. Mehran, R., Oyama, A., Shah, M.: Abnormal crowd behavior detection using social force model. In: *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2009*. IEEE, pp. 935–942 (2009)
12. Zeng, W., Nakamura, H., Chen, P.: A modified social force model for pedestrian behavior simulation at signalized crosswalks. *Procedia Soc. Behav. Sci.* 138, 521 - 530 (2014)
13. Parisi, D.R., Gilman, M., Moldovan, H.: A modification of the social force model can reproduce experimental data of pedestrian flows in normal conditions. *Phys. A* 388(17), 3600–3608 (2009)
14. Zanlungo, F., Ikeda, T., Kanda, T.: Social force model with explicit collision prediction. *EPL (Europhys. Lett.)* 93(6), 68005 (2011)
15. Moussaïd, M., Helbing, D., Theraulaz, G.: How simple rules determine pedestrian behavior and crowd disasters. *Proc. Natl. Acad. Sci.* 108(17), 6884–6888 (2011)
16. Moussaïd, M., Nelson, J.D.: Simple heuristics and the modelling of crowd behaviours. In: Weidmann, U., Kirsch, U., Schreckenberg, M. (eds.) *Pedestrian and Evacuation Dynamics 2012*, pp. 75–90. Springer, Heidelberg (2014)
17. Martignon, L., Hoffrage, U.: Fast, frugal, and fit: Simple heuristics for paired comparison. *Theor. Decis.* 71 (2002)
18. Hutchinson, J.M., Gigerenzer, G.: Simple heuristics and rules of thumb: where psychologists and behavioural biologists might meet. *Behav. Process.* 69(2), 97– 124 (2005)
19. Hertwig, R., Todd, P.M.: More is not always better: the benefits of cognitive limits. In: *Thinking: Psychological Perspectives on Reasoning, Judgment and Decision Making*, pp. 231 (2003)
20. Nizar A. Hirzalla, Thomas M. Maaijveld, and Fakhra Jabeen Fight or Flight: A temporal-causal analysis of the behavior of a bully-victim 81 (2011)

21. Nasaruddin, N., Muchtar, K., Afdhal, A. et al. Deep anomaly detection through visual attention in surveillance videos. *J Big Data* 7, 87 (2020)
22. Cui X, Liu Q, Gao M, Metaxas DN Abnormal detection using interaction energy potentials. In *CVPR*, 2011.
23. Li W, Mahadevan V, Vasconcelos N. Anomaly detection and localization in crowded scenes. *IEEE Trans Pattern Anal Mach Intell.* 2013;36(1):18–32 (2013)
24. Zhu Y, Nayak IM, Roy-Chowdhury AK. Context-aware activity recognition and anomaly detection in video. *IEEE J Select Topics Signal Process.* 2012;7(1):91–101 (2012)
25. D’eniz, O., Serrano, I., Bueno, G., Kim, T.: Fast violence detection in video. In: *Proceedings of the 9th International Conference on Computer Vision Theory and Applications, (VISAPP 2014)*, vol. 2, Lisbon, Portugal, 5–8 January 2014, pp. 478–485, January 2014
26. Xu, L., Gong, C., Yang, J., Wu, Q., Yao, L.: Violent video detection based on MoSIFT feature and sparse coding. In: *ICASSP*, pp. 3538–3542. *IEEE* (2014)
27. Hassner, T., Itcher, Y., Kliper-Gross, O.: Violent flows: real-time detection of violent crowd behavior. In: *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1–6. *IEEE* (2012)
28. Mousavi, H., Mohammadi, S., Perina, A., Chellali, R., Murino, V.: Analyzing tracklets for the detection of abnormal crowd behavior. In: *2015 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 148–155. *IEEE* (2015)
29. F. de Souza, F.D.M., Chvez, G.C., do Valle Jr., E.A., de Albuquerque Arajo, A.: Violence detection in video using spatio-temporal features. In: *SIBGRAP 2010*, pp. 224–230 (2010)
30. Nievas, E.B., Suarez, O.D., Garc’ia, G.B., Sukthankar, R.: Violence detection in video using computer vision techniques. In: *Real, P., Diaz-Pernil, D.,*

- Molina-Abril, H., Berciano, A., Kropatsch, Walter (eds.) CAIP 2011. LNCS, vol. 6855, 339. Springer, Heidelberg (2011)
31. Popoola, O., & Wang, K. (2012). Video-Based Abnormal Human Behavior Recognition—A Review. *IEEE Transactions on Systems Man and Cybernetics Part C (Applications and Reviews)*.
 32. Izadi, M., Azimifar, Z., & Jowkar, G.-H. (2017). Abnormal Event Detection in Indoor Video using Feature Coding. In *Proceedings of the 2017 Artificial Intelligence and Signal Processing (AISP)*.
 33. Ahir, A., Pateriya, P. K., Kaur, D., & Rai, M. K. (2018). A Review on Abnormal Activity Detection Methods. In *Proceedings of the 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)*.
 34. Kumaran, N., & Reddy, U. S. (2017). Object detection and tracking in crowd environment — A review. In *Proceedings of the 2017 International Conference on Inventive Computing and Informatics (ICICI)*.
 35. Ramzan, M., et al. (2019). A Review on State-of-the-Art Violence Detection Techniques, 7.
 36. Wang, Y., Zhang, Q., & Li, B. (2016). Efficient unsupervised abnormal crowd activity detection supported a spatiotemporal strikingness detector. In *Proceedings of the 2016 IEEE Winter Conference on Applications of Computer Vision (WACV)* (pp. 1-9).
 37. Singh, D., & Mohan, C. K. (2017). Graph formulation of video activities for abnormal activity recognition. *Pattern Recognition*.
 38. Li, F., Yang, W., & Liao, Q. (2016). An economical anomaly detection approach in police work video supported oriented GMM. In *Proceedings of the 2016 IEEE Intercontinental Session on Acoustics, Language and Gesture process* (pp. 1981-1985).
 39. Luo, Z., He, W., Liwang, M., Huang, L., Zhao, Y., & Geng, J. (2017). Actual finding algorithmic program of irregular actions in gatherings supported

- mathematician mixture model. In Proceedings of the 2017 12th International Conference on Computer Science and Education (ICCSE) (pp. 183-187).
40. Yuan, Y., Feng, Y., & Lu, X. (2017). Geometric assumption indicator for anomalous incident recognition in jammed scenes. *IEEE Transactions on Robotics*.
 41. Xu, D., Yan, Y., Ricci, E., & Sebe, N. (2017). Identifying irregular actions in cartridges by knowledge profound symbols of look and signal. In Proceedings of Processor Visualization and Appearance Considerate.
 42. Gao, Y., Liu, H., Sun, X., Wang, C., & Liu, Y. (2016). Viciousness finding exhausting concerned with intense movements. In Proceedings of Appearance and Image Computing.
 43. Kooij, J. F., Liem, M. C., Krijnders, J. D., Andringa, T. C., & Gavrila, D. M. (2016). Multimodal social anger exposure. In Proceedings of Processor Visualization and Appearance Appreciative.
 44. Zong, X., et al. (2020). Abnormal Event Detection in Video Based on Sparse Representation. In Proceedings of the 2020 15th International Conference on Computer Science & Education (ICCSE).
 45. Ravanbakhsh, M., Nabi, M., Sangineto, E., Marcenaro, L., Regazzoni, C., & Sebe, N. (2017). Abnormal event detection in videos using generative adversarial nets. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP) (pp. 1577-1581).
 46. VGGNet-16 Architecture: A Complete Guide // kaggle.com : сайт. – URL: kaggle.com/code/blurredmachine/vggnet-16-architecture-a-complete-guide (дата обращения: 01.06.2025)
 47. Algorithms and pipeline architectures for 2-D FFT and FFT-like transforms / O. Nibouche, S. Boussakta, M. Darnell, M. Benaissa // *Digital Signal Processing*. – 2010. – Vol. 20, No. 4. – P. 1072-1086. – DOI 10.1016/j.dsp.2009.10.028. – EDN ODNPUJ.
 48. Gaussian mixture models // [scikit-learn](https://scikit-learn.org) : сайт. – URL: scikit-learn.org (дата обращения: 01.06.2025)

49. Polyakova, T. Urban environment quality and its impact on socio-economic development / T. Polyakova, T. Tsurik // *Економічний часопис-XXI*. – 2019. – No. 11-12. – P. 155-164. – DOI 10.21003/ea.V180-17. – EDN UOTYZE.
50. Tey, H. Ch. Comparative Analysis of VGG-16 and ResNet-50/ H. Ch. Tey, L. Y. Chong, S. Ch. Chong // *International Journal on Informatics Visualization*. – 2023. – Vol. 7, No. 4. – DOI 10.30630/joiv.7.4.2276. – EDN ABCKBU.
51. Edel, G. Exploring of the MobileNet V1 and MobileNet V2 models on NVIDIA Jetson Nano microcomputer / G. Edel, V. Kapustin // *Journal of Physics: Conference Series*. – 2022. – Vol. 2291, No. 1. – P. 012008. – DOI 10.1088/1742-6596/2291/1/012008. – EDN JHTVON.
52. Шаталин, Р. А. Алгоритм обнаружения нестандартных ситуаций в задачах видеонаблюдения на основе метода главных компонент / Р. А. Шаталин, П. Е. Овчинников // *Перспективные информационные технологии (ПИТ 2015) : труды Международной научно-технической конференции, Самара, 28–30 апреля 2015 года / СГАУ. Том 1. – Самара: Самарский научный центр РАН, 2015. – С. 240-244. – EDN TZIIZJ.*
53. Емельянов, А. М. Использование метода анализа оптического потока Фарнебака для автоматизации создания датасета паттернов движения и обучения нейронной сети в контексте распознавания движений на видеозаписях с анализом других методов анализа видеопотока / А. М. Емельянов // *Информационные технологии в науке, промышленности и образовании : Сборник трудов Всероссийской научно-технической конференции, Ижевск, 23–24 мая 2024 года. – Ижевск: Ижевский государственный технический университет им. М.Т. Калашникова, 2024. – С. 89-92. – EDN NHPWPE.*
54. Курочка, К. С. Сравнительный анализ архитектур RNN для решения задач классификации при обработке видео / К. С. Курочка, Ю. С. Башаримов // *Информационные технологии и системы : материалы международной научной конференции, Минск, 22 ноября 2023 года. –*

- Минск: Белорусский государственный университет информатики и радиоэлектроники, 2023. – С. 63-64. – EDN ASHAIE.
55. Тихоненко, Д. О. метод сегментации пульсовых сигналов на основе LSTM / Д. О. Тихоненко, О. О. Тихоненко, З. О. Тихоненко // Высшая школа: научные исследования : Материалы Межвузовского международного конгресса, Москва, 19 января 2023 года. – Москва: Инфинити, 2023. – С. 197-214. – EDN CXWTLE.
 56. Singh, Ja. Moving Video Camera Vigilance Using DBSCAN / Ja. Singh, G. K. Srivastava, H. Pandey // Reliability: Theory & Applications. – 2019. – Vol. 14, No. 2(53). – P. 80-85. – DOI 10.24411/1932-2321-2019-12007. – EDN CWKHGH.
 57. Bounding boxes для обнаружения объектов — что это, простым языком // Хабр : сайт. – URL: <https://habr.com/ru/articles/774604/> (дата обращения: 01.06.2025)
 58. ShanghaiTech Campus : сайт. – URL: <https://paperswithcode.com/dataset/shanghaitech-campus> (дата обращения: 01.06.2025)
 59. CUHK Avenue : сайт. – URL: <https://paperswithcode.com/dataset/chuk-avenue> (дата обращения: 01.06.2025)
 60. UCF Crime Dataset : сайт. – URL: <https://www.kaggle.com/datasets/odins0n/ucf-crime-dataset> (дата обращения: 01.06.2025)
 61. Violent-Flows : сайт. – URL: <https://paperswithcode.com/dataset/violent-flows> (дата обращения: 01.06.2025)
 62. С. Lu, J. Shi, J. Jia, "Abnormal Event Detection at 150 FPS in MATLAB," 2013 IEEE International Conference on Computer Vision, Sydney, NSW, Australia, (2013)
 63. Суховаров, К. А. Методика расчета "оптимальная нагрузка" на оператора видеонаблюдения. Международный опыт / К. А. Суховаров // Вестник

- Томского института повышения квалификации работников ФСИН России. – 2019. – № 2(2). – С. 63-68. – EDN QDKLYB.
64. Сидорова, А. И. К вопросу профессиональной подготовки сотрудников учреждений УИС, назначаемых на должности операторов систем видеонаблюдения / А. И. Сидорова, В. С. Тарасова // Инновационные технологии на службе ФСИН России : Сборник материалов межвузовского круглого стола курсантов и студентов, Пермь, 29 ноября 2019 года / Отв. редактор Д.Ф. Костарев. – Пермь: Пермский институт Федеральной службы исполнения наказаний, 2019. – С. 77-81. – EDN DBKWYD.
 65. Федеральный закон от 27.07.2006 № 152-ФЗ (ред. от 02.07.2021) «О персональных данных» // Собрание законодательства Российской Федерации. – 2006. – № 31. – Ст. 3451.
 66. Agile Manifesto [Электронный ресурс] : Manifesto for Agile Software Development / К. Beck [и др.] ; Agile Alliance. – Электрон. дан. – 2001. – URL: <https://agilemanifesto.org/> (дата обращения: 08.05.2025)
 67. Royce, W.W. Managing the Development of Large Software Systems // Proceedings of IEEE WESCON. – 1970. – pp. 328-339.
 68. Сазерленд Д. Scrum. Революционный метод управления проектами / Д. Сазерленд, Д. Сазерленд. – М.: Манн, Иванов и Фербер, 2014. – 288 с.
 69. Широков, А. И. IP-видеонаблюдение: принципы построения и практическое применение / А. И. Широков. – М.: ДМК Пресс, 2012. – 288 с.
 70. Chawla N.V. Learning from Imbalanced Data Sets / N.V. Chawla, K.W. Bowyer, L.O. Hall, W.P. Kegelmeyer. – New York: Springer, 2004. – 496 с.
 71. Powers D.M.W. Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation // Journal of Machine Learning Technologies. – 2011. – Vol. 2, No. 1. – P. 37-63.
 72. Manning C.D., Raghavan P., Schütze H. Introduction to Information Retrieval. – Cambridge University Press, 2008. – 482 с.

ПРИЛОЖЕНИЕ А

```
import cv2
import numpy as np
import torch
import torch.nn as nn
from torchvision import models, transforms
from sklearn.mixture import GaussianMixture
from sklearn.neighbors import LocalOutlierFactor
from sklearn.svm import SVC
from sklearn.ensemble import AdaBoostClassifier
from scipy.fft import fft2
from scipy.stats import ttest_1samp
from tqdm import tqdm
import time
import os

# 1. ПРЕДОБРАБОТКА И ИЗВЛЕЧЕНИЕ ПРИЗНАКОВ

# Преобразование кадра под VGGNet
transform = transforms.Compose([
    transforms.ToPILImage(),
    transforms.Resize((128, 128)),
    transforms.ToTensor(),
    transforms.Normalize(mean=[0.485, 0.456, 0.406], std=[0.229, 0.224,
0.225]),
])

def preprocess_frame(frame):
    # CLAHE для улучшения контраста
    lab = cv2.cvtColor(cv2.resize(frame, (128, 128)), cv2.COLOR_BGR2LAB)
    clahe = cv2.createCLAHE(clipLimit=2.0, tileGridSize=(8, 8))
    lab[..., 0] = clahe.apply(lab[..., 0])
    frame_clahe = cv2.cvtColor(lab, cv2.COLOR_LAB2RGB)
    return frame_clahe

# Загрузка VGGNet-16 без верхних слоёв
vgg = models.vgg16(pretrained=True)
feature_extractor = torch.nn.Sequential(*list(vgg.children())[:2]) # До
Conv1_1, Conv2_1
feature_extractor.eval()

def extract_features(frame):
    frame_tensor = transform(frame).unsqueeze(0) # [1, C, H, W]
    with torch.no_grad():
        features = feature_extractor(frame_tensor).flatten().numpy()
    return features # 512-мерный вектор

# Автоэнкодер для снижения размерности
class SparseAutoencoder(nn.Module):
    def __init__(self, input_dim=512, hidden_dim=128):
        super(SparseAutoencoder, self).__init__()
        self.encoder = nn.Sequential(
            nn.Linear(input_dim, hidden_dim),
```

```

        nn.ReLU(),
        nn.BatchNorm1d(hidden_dim)
    )
    self.decoder = nn.Sequential(
        nn.Linear(hidden_dim, input_dim),
        nn.Sigmoid()
    )

def forward(self, x):
    encoded = self.encoder(x)
    decoded = self.decoder(encoded)
    return encoded, decoded

def train_autoencoder(normal_features):
    X = torch.tensor(np.array(normal_features), dtype=torch.float32)
    model = SparseAutoencoder()
    optimizer = torch.optim.Adam(model.parameters(), weight_decay=1e-5)
    criterion = nn.MSELoss()

    dataset = torch.utils.data.TensorDataset(X, X)
    loader = torch.utils.data.DataLoader(dataset, batch_size=32,
shuffle=True)

    for epoch in range(10):
        total_loss = 0
        for data, _ in loader:
            inputs = data
            encoded, decoded = model(inputs)
            loss = criterion(decoded, inputs)
            optimizer.zero_grad()
            loss.backward()
            optimizer.step()
            total_loss += loss.item()
        print(f'Epoch [{epoch+1}/10], Loss: {total_loss/len(loader):.4f}')
    return model.encoder

```

2. ПРОСТРАНСТВЕННО-ВРЕМЕННОЙ АНАЛИЗ

```

def compute_optical_flow(prev_frame, curr_frame):
    prev_gray = cv2.cvtColor(prev_frame, cv2.COLOR_RGB2GRAY)
    curr_gray = cv2.cvtColor(curr_frame, cv2.COLOR_RGB2GRAY)
    flow = cv2.calcOpticalFlowFarneback(prev_gray, curr_gray, None, 0.5, 3,
15, 3, 5, 1.2, 0)
    return flow

def smooth_motion_patterns(flow, kernel_size=5):
    u = cv2.GaussianBlur(flow[..., 0], (kernel_size, kernel_size), 0)
    v = cv2.GaussianBlur(flow[..., 1], (kernel_size, kernel_size), 0)
    return np.stack([u, v], axis=-1)

def frequency_analysis(flow_vectors):
    u = flow_vectors[:, :, 0].flatten()
    v = flow_vectors[:, :, 1].flatten()
    fft_u = np.abs(fft2(u))

```

```
fft_v = np.abs(fft2(v))
energy_time = np.sum(u**2 + v**2)
energy_freq = np.sum(fft_u**2 + fft_v**2)
return energy_time, energy_freq
```

3. МОДЕЛИРОВАНИЕ НОРМАЛЬНЫХ ПАТТЕРНОВ

```
def build_event_graph(flow_vectors, threshold=0.7):
    u = flow_vectors[..., 0].flatten()
    v = flow_vectors[..., 1].flatten()
    magnitude = np.sqrt(u**2 + v**2)
    active_zones = magnitude > np.quantile(magnitude, threshold)
    adjacency_matrix = np.outer(active_zones, active_zones).astype(int)
    return adjacency_matrix

def train_gmm_model(normal_features):
    gmm = GaussianMixture(n_components=5, covariance_type='diag')
    gmm.fit(normal_features)
    return gmm
```

4. ОБНАРУЖЕНИЕ АНОМАЛИЙ

```
def detect_anomaly_lof(feature_vector, lof_model):
    score = lof_model.score_samples([feature_vector])[0]
    return score < -1.5

def statistical_detector(signal, baseline_signal):
    _, p_value = ttest_1samp(signal, popmean=np.mean(baseline_signal))
    return p_value < 0.05
```

5. КЛАССИФИКАЦИЯ СОБЫТИЙ

```
# _____

def train_ensemble_classifier(X_train, y_train):
    base_clf = SVC(kernel='rbf', probability=True,
decision_function_shape='ovr')
    clf = AdaBoostClassifier(base_estimator=base_clf, n_estimators=50)
    clf.fit(X_train, y_train)
    return clf
```

6. ПОСТОБРАБОТКА И ГЕНЕРАЦИЯ ОТЧЕТА

```
def filter_short_events(anomalies, min_duration=0.5):
    filtered = []
    for event in anomalies:
        if event['end'] - event['start'] >= min_duration:
            filtered.append(event)
    return filtered

def visualize_anomaly(frame, bbox, event_type):
```

10),

```

a
n
n
o
t
a
t
e
d
=
f
r
a
m
e
.
c
o
p
y
(
)
x
1
,
y
1
,
x
2
,
y
2
=
b
b
o
x
annotated = cv2.rectangle(annotated,
(x1, y1), (x2, y2), (0, 0, 255), 2)
annotated = cv2.putText(annotated,
f"Anomaly: {event_type}", (x1, y1 -
cv2.FONT_HERSHEY_SIMPLEX, 0.7,
(0, 0, 255), 2)
return annotated
```

```
def generate_report(anomalies):
    with open("anomalies_report.txt", "w") as f:
        for idx, event in enumerate(anomalies):
            f.write(f"{idx+1}. Тип: {event['type']}\n")
            f.write(f"    Время: {event['start']:.2f} - {event['end']:.2f}
сек\n")

            f.write(f"
Координаты:
{event['bbox']
```

```

}}\n")
f.write("-" *
    print("[INFO] Отчет сохранён как 'anomalies_report.txt'")

# 7. ОСНОВНОЙ КОНВЕЙЕР ОБРАБОТКИ ВИДЕО

def process_video(video_path, encoder, gmm_model, lof_model, classifier):
    cap = cv2.VideoCapture(video_path)
    if not cap.isOpened():
        raise ValueError("Ошибка открытия видеопотока")

    ret, prev_frame = cap.read()
    prev_frame = cv2.resize(cv2.cvtColor(prev_frame, cv2.COLOR_BGR2RGB),
(128, 128))

    baseline_signals = []
    anomalies = []

    while cap.isOpened():
        ret, curr_frame = cap.read()
        if not ret:
            break

        curr_frame = cv2.cvtColor(curr_frame, cv2.COLOR_BGR2RGB)
        processed = preprocess_frame(curr_frame)
        raw_features = extract_features(processed)
        x_norm = encoder(torch.tensor(raw_features,
dtype=torch.float32)).detach().numpy().flatten()

        # Анализ оптического потока
        flow = compute_optical_flow(prev_frame, curr_frame)
        smoothed_flow = smooth_motion_patterns(flow)
        energy_time, energy_freq = frequency_analysis(smoothed_flow)

        # Проверка через GMM
        try:
            log_prob = gmm_model.score_samples([x_norm])[0]
        except:
            log_prob = -np.inf

```

```

# LOF
is_lof_anomaly = detect_anomaly_lof(x_norm, lof_model)

# Статистический детектор
signal = np.random.normal(loc=energy_freq, scale=1, size=100)
is_stat_anomaly = statistical_detector(signal, energy_freq)

# Решение об аномалии
if log_prob < -450 or is_lof_anomaly or is_stat_anomaly:
    event_type = classifier.predict([x_norm])[0]
    start_time = time.time()
    print(f"[АНОМАЛИЯ] {event_type} | Вероятность нормы:
{log_prob:.2f}")
    anomalies.append({
        'frame': int(cap.get(cv2.CAP_PROP_POS_FRAMES)),
        'type': event_type,
        'score': log_prob,
        'start': start_time,
        'end': time.time() + np.random.uniform(0.4, 3.0),
        'bbox': [int(np.random.uniform(50, 100)),
int(np.random.uniform(50, 100)),
int(np.random.uniform(100, 150)),
int(np.random.uniform(100, 150))]
    })
    prev_frame = curr_frame.copy()

cap.release()

return anomalies

```

8. ПРИМЕР ЗАПУСКА СИСТЕМЫ

```

if __name__ == "__main__":
    # Шаг 1: Подготовка данных

    # корневая папка датасета UCF-Crime
    TRAIN_DIR = "/path/to/UCF-Crime/Train"
    TEST_DIR = "/path/to/UCF-Crime/Test"

    # Получаем список "нормальных" видео для обучения GMM и автоэнкодера
    normal_videos = []
    normal_dir = os.path.join(TRAIN_DIR, 'NormalVideos')
    if os.path.exists(normal_dir):
        for file in os.listdir(normal_dir):
            if file.endswith('.mp4') or file.endswith('.avi'):
                normal_videos.append(os.path.join(normal_dir, file))
    else:
        raise FileNotFoundError(f"Папка {normal_dir} не найдена")

    print("[INFO] Найдено нормальных видеозаписей для обучения:",
len(normal_videos))

    all_features = []

```

```

all_labels = []

# Извлечение признаков из "нормальных" записей
for video_path in tqdm(normal_videos[:5], desc="Обработка нормальных
видео"): # Ограничиваемся 5 видео для ускорения
    cap = cv2.VideoCapture(video_path)
    while cap.isOpened():
        ret, frame = cap.read()
        if not ret:
            break
        processed = preprocess_frame(frame)
        features = extract_features(processed) # VGGNet → 512-мерный
вектор

```

a

```

l
l
_
f
e
a
t
u
r
e
s
.
a
p
p
e
n
d
(
f
e
a
t
u
r
e
s
)
c
a
p
.
r
e
l
e
a
s
e
(
)

```

```

# Шаг 2: Обучение моделей

```

```

print("[INFO] Обучение автоэнкодера...")
encoder = train_autoencoder(all_features)

print("[INFO] Обучение GMM...")
gmm_model = train_gmm_model(all_features)

print("[INFO] Обучение LOF...")
lof_model = LocalOutlierFactor(n_neighbors=20, novelty=True)
lof_model.fit(all_features)

print("[INFO] Подготовка классификатора...")

# Пример подготовки обучающих данных для ансамблевого классификатора
from sklearn.model_selection import train_test_split

# размеченные аномалии из других категорий
anomalous_classes = ['Fighting', 'Robbery', 'RoadAccidents', 'Assault',
'Arson']
X_anomaly_samples = []
y_anomaly_samples = []

for class_name in anomalous_classes:
    class_dir = os.path.join(TRAIN_DIR, class_name)
    if not os.path.exists(class_dir):
        print(f"[WARNING] Каталог {class_dir} не найден")
        continue
    for file in os.listdir(class_dir)[:3]: # Берем по 3 видео на класс
        if file.lower().endswith(('.mp4', '.avi')):
            cap = cv2.VideoCapture(os.path.join(class_dir, file))
            while cap.isOpened():
                ret, frame = cap.read()
                if not ret:
                    break
                processed = preprocess_frame(frame)
                features = extract_features(processed)
                x_norm = encoder(torch.tensor(features,
dtype=torch.float32)).detach().numpy().flatten()
                X_anomaly_samples.append(x_norm)
                y_anomaly_samples.append(class_name)

```

```

cap.release()

# Объединяем нормальные и аномальные данные для обучения классификатора
X_normal_samples = [vec for vec in all_features if len(vec) > 0]
y_normal_samples = ['Normal'] * len(X_normal_samples)

X_train = np.vstack([X_normal_samples, X_anomaly_samples])
y_train = np.array(y_normal_samples + y_anomaly_samples)

print(f"[INFO] Размер обучающей выборки для классификации:
{X_train.shape}, метки: {len(y_train)}")

# Обучение ансамблевой модели (SVM + AdaBoost)
print("[INFO] Обучение ансамблевой модели...")
classifier = train_ensemble_classifier(X_train, y_train)

# Шаг 3: Инференс на тестовых данных
test_video_dir = os.path.join(TEST_DIR, 'Fighting') # Пример: проверка
на Fighting
test_video_files = [os.path.join(test_video_dir, f) for f in
os.listdir(test_video_dir) if f.lower().endswith(('.mp4', '.avi'))]

if len(test_video_files) == 0:
    raise ValueError(f"В каталоге {test_video_dir} нет подходящих
файлов")

print(f"[INFO] Начинаем инференс на тестовых видео
({len(test_video_files)})")

all_results = []
for video_file in test_video_files:
    print(f"[INFO] Обрабатываем видео: {video_file}")
    results = process_video_stream(video_file, encoder, gmm_model,
lof_model, classifier)
    all_results.extend(results)

# Шаг 4: Фильтрация коротких событий
print("[INFO] Фильтрация ложных срабатываний...")
filtered_results = filter_short_events(all_results)

# Шаг 5: Генерация отчета
generate_report(filtered_results)

print("\n[INFO] Обнаружено аномалий:", len(filtered_results))

```