

ТОМСКИЙ
ПОЛИТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ



На правах рукописи



Лаптев Никита Витальевич

**ОБНАРУЖЕНИЕ ОБЪЕКТОВ ВИДЕОРЯДА
НА ОСНОВЕ ТЕХНОЛОГИИ
МАШИННОГО ОБУЧЕНИЯ
(НА ПРИМЕРЕ ЛЕСНЫХ ПОЖАРОВ)**

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

2.3.1 – Системный анализ, управление и обработка информации, статисти-
стика

Томск – 2023

Работа выполнена в федеральном государственном автономном образовательном учреждении высшего образования "Национальный исследовательский Томский политехнический университет"

Научный руководитель:

Гергет Ольга Михайловна

доктор технических наук, доцент, профессор
Национальный исследовательский Томский политехнический университет

Официальные оппоненты:

Ходашинский Илья Александрович

доктор технических наук, профессор, ФГБОУ ВО «Томский государственный университет систем управления и радиоэлектроники», профессор кафедры компьютерных систем в управлении и проектировании

Николайчук Ольга Анатольевна

доктор технических наук, доцент, ФГБНУ «Институт динамики систем и теории управления имени В.М. Матросова СО РАН», старший научный сотрудник лаборатории Методов автоматизации исследования техногенной безопасности

Защита состоится 13.09.2023 г. в 15 часов 00 мин на заседании диссертационного совета ДС.ТПУ.15 при федеральном государственном автономном образовательном учреждении высшего образования "Национальный исследовательский Томский политехнический университет" по адресу: Россия, 634034, г. Томск, ул. Советская, 84/3, ауд. 214.



С диссертацией можно ознакомиться в научно-технической библиотеке Томского политехнического университета и на сайте dis.tpu.ru при помощи QR-кода.

Автореферат разослан "___" _____ 2023 г.

Ученый секретарь

диссертационного совета ДС.ТПУ.15

д.т.н.

A stylized handwritten signature in blue ink, consisting of several fluid, connected strokes.

А.Я. Пак

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность работы. Обнаружение объектов в видеоряде является одной из ключевых проблем исследований в области искусственного интеллекта, которые включены в направление Н1 Стратегии научно-технического развития Российской Федерации. При разработке эффективных методов и алгоритмов обнаружения объектов на основе таких признаков как форма, цвет и текстура требуется учитывать ряд мешающих факторов: размытие объектов при движении, вариативность признаков, изменение условий освещения, перекрытие объектов в процессе перемещения и другие. Игнорирование этих факторов может привести к снижению точности обнаружения и увеличению вероятности ложного срабатывания систем распознавания.

Перспективными инструментами для решения задач анализа статических и динамических признаков на кадрах видеозаписей являются технологии машинного обучения, в частности нейронные сети (НС). Для их применения необходимо осуществлять сбор, разметку и предварительную подготовку исходных данных; выбор модели и подбор параметров архитектуры НС; тонкую настройку гиперпараметров НС, т.е. скорости обучения и объема данных (размера батча) с целью снижения времени сходимости в процессе обучения; интерпретацию результатов анализа данных НС, в том числе анализ хронологии изменения объекта в видеоряде.

Задачи автоматического обнаружения динамических объектов в видеопотоке являются наиболее актуальными в таких приложениях, как автоматическое распознавание лиц, медицинская диагностика, робототехника, транспортные системы и др.

Одним из приложений, имеющих социально важное значение, является разработка систем раннего обнаружения лесных пожаров. Ежегодно в Российской Федерации фиксируется от 10 тысяч до 35 тысяч лесных пожаров, а их общая площадь превышает 1 млн га. Согласно заявлению Федерального агентства лесного хозяйства, в 2022 году общая площадь лесных пожаров в России составила 8,96 млн га, а за последние 3 года в среднем выгорело около 17,8 млн га леса. При этом экономический ущерб составил в среднем 9,7 млрд рублей.

На точность обнаружения пожаров влияет множество факторов, включая освещенность, погодные условия, прозрачность окружающей среды, схожесть характерных признаков с облачностью и туманностью. Дополнительные сложности при обнаружении пожаров создает непостоянство таких признаков, как форма, цвет и текстура. При этом, пожар является одним из наиболее сложных объектов обнаружения, в силу необходимости разработки специальных методов и алгоритмов, учитывающих мешающие факторы, влияющие на точность и надежность обнаружения объектов в видеоряде.

Целью диссертационной работы является разработка технологии и алгоритмов обнаружения объектов, обладающих динамическими признаками на изображениях видеоряда, обеспечивающих высокое качество распознавания.

В связи с поставленной целью должны быть решены следующие **задачи**:

- провести анализ существующих методов и алгоритмов обнаружения объектов видеоряда с динамическими признаками;

- разработать технологию автоматического обнаружения объектов с динамическими признаками, состоящими из предложенных и реализованных алгоритмов;
- разработать программное обеспечение и провести вычислительный эксперимент с целью оценки качества результатов обнаружения объектов, полученных на основе предлагаемых алгоритмов.

Методы исследования. Использованы методы цифровой обработки изображений, технология компьютерного зрения, методы преобразования и анализа изображений, машинного обучения и прикладного программирования. Экспериментальные численные исследования проводились с использованием методов математического анализа, моделирования и математической статистики.

Научная новизна

- Предложена технология анализа изображений в видеопоследовательности, позволяющая достичь высокой точности обнаружения объектов за счет локализации и классификации каждого объекта с учетом как статических, так и динамических признаков.
- Предложен и обоснован алгоритм объединения предсказаний нейросети, позволяющий объединять плотные группы объектов, выделенные на основе статических признаков в один объект на основе оценки отношения площадей пересечения соседних объектов, используемые для формирования кластеров, локализуящих распознаваемые объекты.
- Разработан алгоритм классификации обнаруженных объектов на основе состоящей из сверточной и рекуррентной нейросетей гибридной архитектуры, обеспечивающей неявное представление хронологии изменения объектов при повышении точности классификации изображений в видеопотоке.

Достоверность полученных результатов диссертационной работы подтверждается результатами тестирования, апробации и внедрения разработанных алгоритмов и технологии на собственных и общедоступных базах изображений.

Практическая ценность работы. Разработанные в диссертационной работе алгоритмические и программные средства могут быть использованы для решения задач обнаружения объектов видеоряда в различных прикладных областях: медицинская диагностика, контроль перемещений роботов, мониторинг местности и т.п.

Положения, выносимые на защиту

1. Технология анализа изображений в видеопоследовательности на основе предложенных алгоритмов локализации и классификации с учетом статических и динамических данных позволяет достичь высокой точности классификации объектов по метрике F-score не менее 95 %.
2. Алгоритм объединения предсказаний нейросети, позволяющий объединять плотные группы объектов, выделенные на основе статических признаков в один объект с целью повышения точности локализации распознаваемых объектов. Реализация предложенного алгоритма позволяет повысить точность локализации на 6 %.

3. Алгоритм классификации обнаруженных объектов на основе предложенной гибридной архитектуры нейросети позволяет увеличить качество классификации по метрике F-score на 9 %.

Реализация и внедрение результатов работы. Результаты диссертационного исследования использованы при выполнении следующих научно-исследовательских проектов: государственное задание "Наука" №FFSWW-2020-0014 "Разработка научных основ технологии роботизированной мультипараметрической томографии на основе методов обработки больших данных и машинного обучения для исследования перспективных композиционных материалов", 2020-2022 гг.; грант РФФИ № 20-37-90055 "Методы и интеллектуальные технологии планирования движения беспилотных транспортных средств", 2020-2022 гг.; грант УМНИК №16068ГУ/2020 "Система детектирования и локализации очага возгорания в видеоряде на основе средств компьютерного зрения", 2020-2023 гг.

Результаты работы также используются: в ООО "НТП КИБЕРЦЕНТР" (г. Томск) – в программном обеспечении информационно-телекоммуникационной системы интеллектуального видеонаблюдения и аналитики лесопожарной обстановки (ИТС «ИВА»), в ООО "Спотпаркинг" (г. Томск) – в составе системы для обнаружения автомобилей на парковочном пространстве и в учебном процессе в отделении информационных технологий Инженерной школы информационных технологий и робототехники ТПУ. Акты о внедрении приложены к диссертационной работе.

Апробация. Основные результаты данной работы освещались на следующих конференциях и симпозиумах: Международной научно-практической конференции студентов, аспирантов и молодых ученых, г. Томск, 2019 г.; 14-й Международный форум по стратегическим технологиям "IFOST 2019", г. Томск, 2019 г.; XXXIV Международная научная конференция "Математические методы в технике и технологиях", г. Санкт-Петербург, 2021 г.; 4-й Международный семинар по фотограмметрии и методам компьютерного зрения для видеонаблюдения, биометрии и биомедицины, г. Москва, 2021 г.; 31-я Международная конференция по компьютерной графике и зрению "ГрафиКон 2021", г. Нижний Новгород, 2021 г.; International Symposium on Electrical, Electronics and Information Engineering, Online, 2021 г.; 6-я Международная конференция по тестированию программного обеспечения, машинному обучению и комплексному анализу процессов, ТМРА 2021, г. Томск, 2021 г.; SIBCON-2022 международная IEEE-сибирская конференция по управлению и связи, г. Томск, 2022 г.

Публикации. Основные результаты исследований отражены в 13 публикациях: 5 статей в ведущих научных журналах и изданиях, рекомендуемых ВАК, из них две статьи проиндексированы в базах данных Scopus и (или) Web of Science; 1 патент на изобретение, 3 свидетельства о государственной регистрации программ для ЭВМ; 4 статьи в сборниках трудов международных и российских конференций.

Структура и объем диссертации. Диссертационная работа состоит из введения, трех глав, заключения, списка литературы из 133 наименований. Работа содержит 163 страниц основного текста, включая 54 рисунков и 10 таблиц.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы исследования, сформулированы цель и задачи работы, определена научная новизна, а также отражены практическая значимость и положения, выносимые на защиту.

В первой главе "Анализ алгоритмов и методов обнаружения объектов" исследованы методы распознавания огня и дыма на изображениях, полученных с оптической камеры. Выявлены существующие проблемы обнаружения с использованием алгоритмов на основе цвета и выделения динамических признаков.

Проведен обзор методов обнаружения объектов на кадрах видеоряда, описаны известные современные нейросетевые методы обнаружения объектов. Подробно рассмотрены модели рекуррентных нейронных сетей, позволяющих решить задачу выделения временной составляющей движения дымового облака. Проведен анализ алгоритмов объединения прогнозов нейросетевого анализа данных, назначение которых выделить исследуемый объект с нескольких кадров. Обоснована необходимость разработки комплексной технологии раннего обнаружения пожаров.

На основе литературного обзора, сделан вывод о том, что для решения поставленной задачи наиболее эффективным является реализация алгоритмов, основанных на ансамблировании методов машинного обучения.

Во второй главе "Технология обнаружения объектов видеоряда на основе технологии машинного обучения (на примере лесных пожаров)" описаны предлагаемые алгоритмы, которые позволяют осуществить классификацию и локализацию объектов исследования в видеоряде.

Разработана технология обнаружения пожароопасных объектов в лесном массиве, которая позволяет анализировать данные, полученные со стационарных видеокамер. Предложенную технологию можно укрупненно представить в виде трех блоков (рисунок 1).

1. Предварительная обработка изображений:

- чтение видеозаписи с последующим разбиением на кадры;
- формирование кадров для анализа с помощью алгоритма выделения динамических признаков.

2. Выделение на кадрах объектов анализа:

- обнаружения объектов исследования в видеоряде и заключение их в ограничивающую рамку;
- объединение разрозненных объектов анализа в один объект на результирующем кадре, посредством использования алгоритма кластеризации площадей обнаруженных объектов.

3. Фильтрация обнаруженных объектов:

- подготовка данных для классификации обнаруженных объектов;
- классификация с помощью алгоритма фильтрации ранее обнаруженных объектов на основе гибридной архитектуры нейронной сети;
- визуализация результатов.

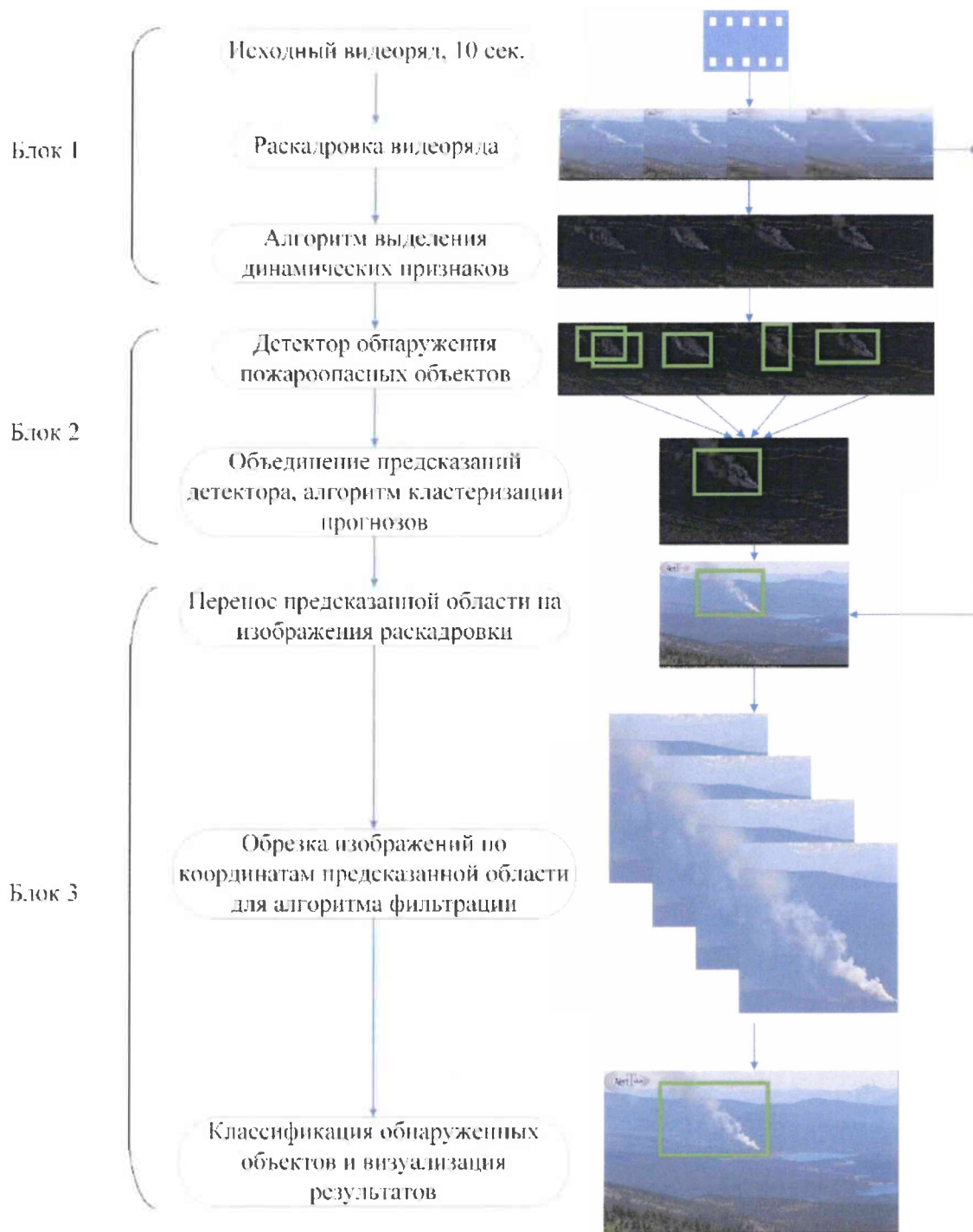


Рисунок 1 – Схема технологии раннего обнаружения пожаров в лесном массиве

Алгоритм выделения динамических признаков. Для выделения динамических признаков на изображении реализован метод разности кадров, преимуществом которого является нечувствительность к изменениям сцены (например, к освещению) и возможность адаптации к различным динамическим средам с хорошей стабильностью. Однако, серьезным недостатком является невозможность извлечения всей площади объекта. В связи с этим при реализации алгоритма происходит вычисление средней интенсивности пикселей кадров последовательности с последующим

вычислением разности с одним из исходных кадров. Последовательность работы алгоритма представлена на рисунке 2 и состоит из следующих этапов:

1. Последовательность кадров делится на n равных интервалов, из которых выбираются значения верхней границы интервала;
2. Операция преобразования кадров из трех каналов RGB в один канал HBR (переход в градации серого), что позволяет экономить время вычислений на следующих этапах.
3. Операция инициализации "усредненного кадра".

$$F = \{F(x, y), 0 \leq x < w, 0 \leq y < h\}, \quad (1)$$

$$F_c(x, y) = \frac{1}{N} \sum_{i=1}^N F_i(x, y), \quad (2)$$

где x – количество пикселей, соответствующих ширине кадра, y – количество пикселей, соответствующих высоте кадра, w – максимальная ширина кадра, h – максимальная высота кадра $F_c(x, y)$ – "усредненный кадр", N – общее число, обрабатываемых кадров. $F_i(x, y)$ – текущий кадр последовательности.

Данная операция позволяет нивелировать проблему шумности камер и получать более стабильный результат.

4. Операция выделения динамических признаков на основе вычитания исходного кадра из усредненного:

$$F_{di}(x, y) = |F_c(x, y) - F_i(x, y)|, \quad i = 1, 2, \dots, N, \quad (3)$$

где $F_{di}(x, y)$ – разность текущего кадра последовательности и усредненного.

5. Операция шумоподавления. Данная операция позволяет выделить объекты с большей динамикой, отбросив при этом посторонние шумы. Фильтрация изображения с адаптивным порогом:

$$F_{ri}(x, y) = \begin{cases} m, & \text{если } F_{di}(x, y) > th \\ 0, & \text{если нет} \end{cases}, \quad (4)$$

где $F_{ri}(x, y)$ – результирующий кадр операции шумоподавления, th – пороговое значение яркости, m – заданное значение яркости



a)



б)

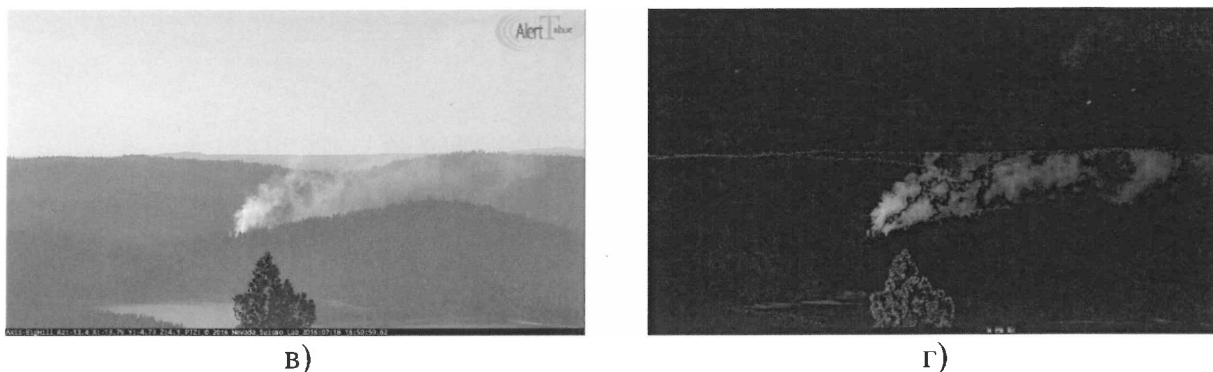


Рисунок 2 – Визуальное представление работы алгоритма выделения динамических признаков: а) исходное изображение; б) перевод изображения из RGB в HBR; в) операция шумоподавления; г) операция вычитания фона.

Алгоритм кластеризации прогнозов обнаруженных объектов (АКП). Интенсивность истечения дыма зависит от нескольких факторов, например, длительности и количества вещества, подвергшегося горению. В результате может возникнуть ситуация, связанная с отсутствием дыма на изображениях в разные моменты времени, поэтому для минимизации ошибки не обнаружения очага возгорания, на вход системы подается серия изображений: соответствующие началу, середине и концу видеоряда. Входное изображение преобразуется к разрешению 512×512 пикселей и имеет выделенные участки, полученные с помощью алгоритма выделения динамических признаков. На выходе системы формируются предсказания объектов: класс, вероятность обнаружения и координаты объектов. В итоге имеем несколько объектов обнаружения с разной вероятностью и разными координатами. Необходимо объединить детекции и усреднить значения вероятностей, если они относятся к одному объекту. На вход алгоритма подается список всех обнаруженных ограничивающих рамок (боксов) на кадре, а также их вероятности принадлежности к классу. Блок-схема алгоритма кластеризации прогнозов обнаруженных объектов представлена на рисунке 3 и состоит из следующих этапов:

1. Фильтрация данных. Проверка на условие прохождения минимального порога по вероятности ограничивающей рамки. Итоговый список содержит в себе все обнаруженные объекты, отсортированные в порядке уменьшения степени вероятности принадлежности к классу.

2. Расчет пересечения боксов по формуле:

$$\Delta S = [\min(p_{X_2}, p_{X_4}) - \max(p_{X_1}, p_{X_3})] \cdot [\min(p_{Y_2}, p_{Y_4}) - \max(p_{Y_1}, p_{Y_2})], \quad (1)$$

где ΔS – площадь пересечения боксов, $p_{(X,Y)_1}$ – координата нижнего левого угла первого прямоугольника, $p_{(X,Y)_2}$ – координата верхнего правого угла первого прямоугольника, $p_{(X,Y)_3}$ – координата нижнего левого угла второго прямоугольника, $p_{(X,Y)_4}$ – координата верхнего правого угла второго прямоугольника.

3. Формирование $n \times n$ матрицы $A = [a_{ij}]$, где n – количество боксов, которые остались после фильтрации. Элементы матрицы заполняются по формуле:

$$a_{ij} = \begin{cases} \frac{\Delta S}{S_i} \cdot 100, \text{ если } i > j \\ -2, \text{ если } i = j \\ \frac{\Delta S}{S_j} \cdot 100, \text{ если } i < j \end{cases}, \quad (2)$$

где S_i – площадь i -го прямоугольника.

4. Заполнение главной диагонали матрицы A номерами кластеров боксов. Кластер y_c с номером c формируется, если имеет место минимальное необходимое количество ms пересекающихся боксов, и элементу a_{ii} присваивается текущий номер кластера. Рекурсивно каждому элементу найденного кластера присваивается номер текущего кластера. Процесс продолжается до тех пор, пока всем диагональным элементам матрицы A не будут присвоены номера кластеров.

5. Расчет центра масс верхнего и нижнего углов ограничивающей рамки для каждого кластера по формуле:

$$P_{(X,Y)_c} = \frac{\sum_i P_{(X,Y)_i} P_i}{\sum_i P_i}, \quad (5)$$

где $P_{(X,Y)_c}$ – координата нового центра угла, $P_{(X,Y)_i}$ – координата угла обнаруженного объекта, P_i – вероятность обнаруженного объекта.

6. Вычисление новой вероятности P_c прогнозируемых боксов в зависимости от количества обнаруженных кластеров по формуле:

$$P_c = \frac{\sum_i P_i^2}{\sum_i P_i}. \quad (6)$$

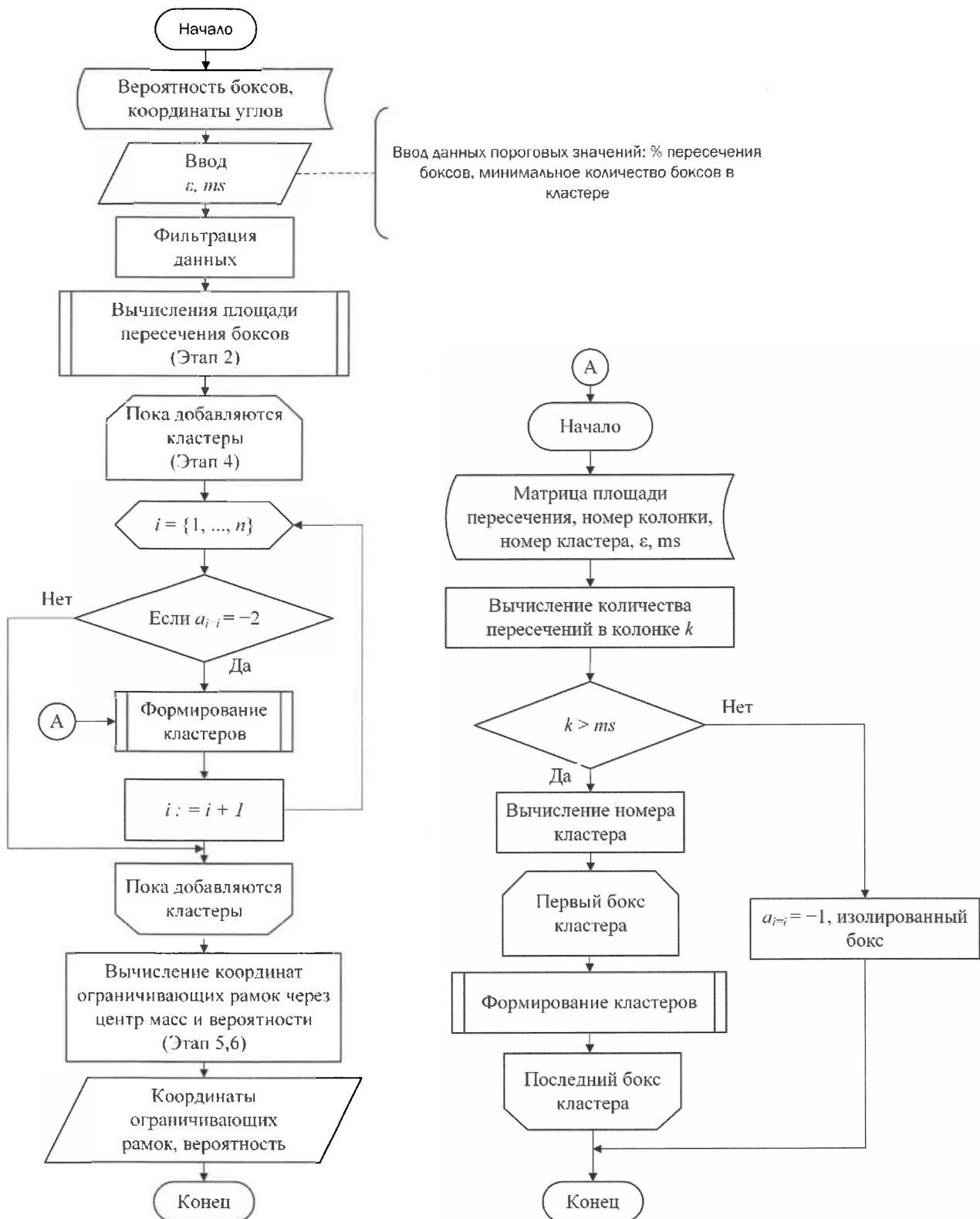


Рисунок 3 – Блок схема АКП обнаруженных объектов

Отличительной особенностью от существующих алгоритмов кластеризации является правило расчета метрики расстояния между точками. Под расстоянием понимается величина равная отношению площади i -го и j -го боксов к площади пересечения боксов (формула 2), таким образом, формируется ассиметричная матрица.

Реализация алгоритма в отличие от существующих позволяет решить проблему кластеризации для асимметричной матрицы расстояний.

Алгоритм фильтрации обнаруженных объектов. Изучаемые видеопотоки содержат различные типы объектов, в частности облака и туманность, которые очень похожи по текстуре на дым от лесного пожара. Следовательно, чтобы правильно классифицировать обнаруженные объекты, необходимо учитывать временное изменение дымового облака. С этой целью сначала выделяются участки изображения с высокой вероятностью обнаружения дыма посредством модели EfficienDet-D0 и алгоритма выделения динамических признаков. Блок-схема алгоритма классификации обнаруженных объектов представлена на рисунке 4. Из кадров вырезаются фрагменты изображений по координатам, полученным посредством работы алгоритма кластеризации прогнозов обнаруженных объектов. Таким образом, формируется последовательность изображений с использованием четырех кадров, как показано на рисунке 1 (блок 3). На основе последовательности фрагментов изображений с областью дыма от лесного пожара, выделяются и классифицируются хронологические изменения объектов с помощью гибридной нейронной сети, включающей сверточную (СНС) и рекуррентную (РНС) нейронные сети.

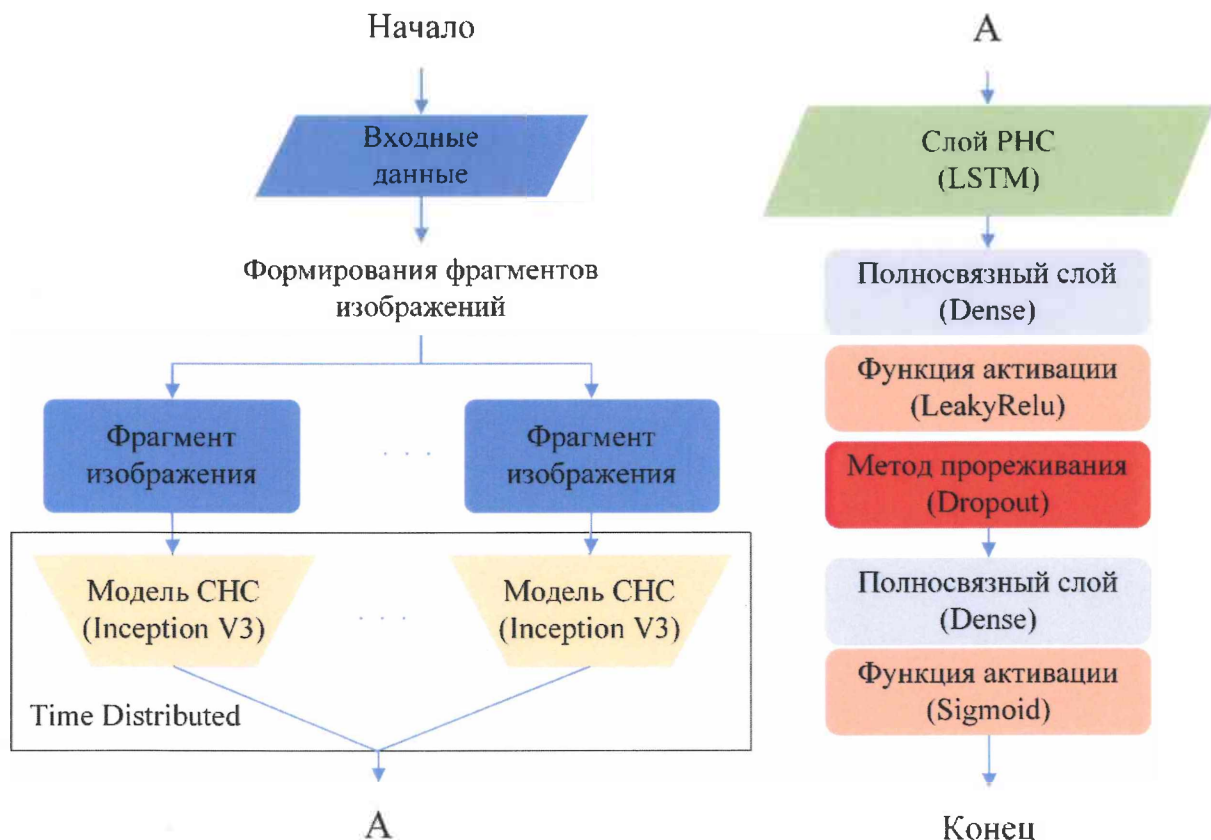


Рисунок 4 – Блок-схема алгоритма классификации обнаруженных объектов, на основе комбинированной архитектуры модели нейронной сети CNN-RNN

Таблица 1 Параметры гибридной архитектуры нейронной сети (выделенные значения, исследуемые параметры)

Тип	Параметры		Характеристика
Слой TimeDistributed (Модель CNN Inception-V3)	Вход	4(300x300x3)	4 изображения
	Выход	4x2048	4 временных шага на вектор признаков
Слой RNN (LSTM)	Вход	4x2048	4 временных шага на вектор признаков
	Выход	1024	Нейронов
Слоя Dense с функцией активации LeakyRelu	Вход	1024	Нейронов
	Выход	256	
Слой Dropout	Вход	256	Коэффициент прореживания 0,5
	Выход	256	
Слоя Dense с функцией активации Sigmoid	Вход	256	Нейронов
	Выход	1	Нейрон с выводом значения вероятности принадлежности к классу от 0 до 1

Принцип работы, представленной архитектуры, заключается в следующем: анализируемые фрагменты кадров видеозаписи передаются модели CNN для извлечения функций низкого уровня. В работе использована модель Inception-V3, предварительно обученная на базе данных изображений ImageNet, в качестве базовой модели CNN. Для формирования итогового набора векторов признаков используется слой TimeDistributed, объединяющий вектора признаков, полученные в ходе анализа каждого изображения. Блок RNN сети состоит из входного слоя, слоя LSTM, полносвязного слоя и слоя Dropout. Выходы модулей LSTM передаются на полносвязный слой с добавлением функции активации LeakyRelu и слоя Dropout. Последний слой сети представлен полносвязным слоем с одним выходом и сигмоидной функцией активации для вывода вероятности классификации, обнаруженных объектов. Ключевой особенностью разработанного алгоритма является неявное представление времени с целью извлечения признаков на основе данных прошлых изображений, хранящихся в памяти НС.

В третьей главе "Тестирование предложенных алгоритмов в рамках технологии обнаружения объектов видеоряда на основе средств машинного обучения" отражены результаты, полученные при реализации алгоритма предобработки кадров для анализа моделью нейронной сети, а также выборе модели нейронной сети для обнаружения области исследования; оптимизации гиперпараметров гибридной архитектуры нейронной сети для классификации обнаруженных объектов.

Исходной информацией для проведенных экспериментов являются данные, полученные из нескольких источников, в частности, данные открытых онлайн ресурсов Nevada Seismological Laboratory университета города Невады (США), Center for wildfire research университета города Сплит (Хорватия), видеозаписи съемки лесных пожаров Пермского и Калининградского лесных хозяйств, видеозаписи базы

данных Калифорнийского университета в Сан-Диего "HPWREN". С целью непосредственного обучения все данные были предварительно размечены с использованием средств web-сервиса "SuperVisely". Общее количество собранных видеозаписей составило 550 единиц, из них содержащих очаги возгорания – 350 единиц и 200 видеозаписи без пожаров.

Предварительная обработка, обучение, тестирование и разметка данных осуществлялась на компьютере со следующими характеристиками: Intel Core i7-4820K CPU 3.7GHz, Nvidia GeForce GTX2080Ti, 32GB RAM и Windows10.

Разработанные алгоритмы внедрены в программно-аппаратный комплекс, реализующий: захват видеопотока с цифровой камеры и запуск библиотеки для анализа данных. В библиотеке реализованы следующие функции: предобработка данных, запуск нейронных сетей для трекинга объектов, постобработка результатов трекинга, обрезка кадров по результатам постобработки, формирование серии изображений, фильтрация обнаруженных объектов, визуализация результатов. Библиотека имеет пользовательский интерфейс для настройки параметров алгоритмов предобработки и постобработки пользователем. Интерфейс программы написан на языках программирования C и C++.

Алгоритм выделения динамических признаков на изображении. Для проверки эффективности алгоритма были обучены 5 зарекомендовавших себя моделей нейронных сетей на разных выборках данных, а именно: изображения *без цифровой обработки* (таблица 2, столбец "БЦО"), изображения *с обработкой с помощью алгоритма выделения динамических признаков без цветовой обработки* (таблица 2, столбец "ВФ"), изображения *с обработкой с помощью алгоритма выделения динамических признаков с удалением шумов фильтром Гаусса* (таблица 2, столбец "ФГ"), изображения *с обработкой с помощью алгоритма выделения динамических признаков с цветовой обработкой, перевод в один канал и удаление шумов фильтром Гаусса* (таблица 2, столбец "НБР"). В таблице 2 приведены значения р-критерия (p-value) и $\bar{X}$ – среднего значения метрики точности и полноты.

Для сравнения результатов обнаружения использовались метрики полноты (Recall) и точности (Precision). На основе проведенного исследования выявлено, что включение этапа предобработки при обнаружении объектов влияет на точность и полноту результатов вне зависимости от архитектуры нейронной сети. Реализованный алгоритм выделения динамических признаков и выбранная модель Efficient-D0 успешно выделяют большинство очагов возгорания на изображениях. При этом полнота составила 89 % при пороге вероятности обнаружения ≥ 35 %. На данном этапе работы системы наиболее ценным является сравнение по метрике полноты, показатели метрики точности будут улучшены в результате фильтрации обнаруженных объектов на следующих этапах работы.

Таблица 2 – Таблица эффективности различных алгоритмов предобработки данных по критерию Манна-Уитни

Метрика	Модель НС	ВФ		ФГ		HBR		БЦО
		p	$\bar{X}$	p	$\bar{X}$	p	$\bar{X}$	$\bar{X}$
Точность	Efficient-D0	0,67	0,70	0,16	0,81	0,09	0,85	0,64
	Efficient-D1	0,75	0,78	0,83	0,75	0,03	0,87	0,68
	Efficient-D2	0,36	0,80	0,15	0,77	0,63	0,80	0,72
	Faster-resnet50	0,02	0,85	0,43	0,81	0,04	0,81	0,76
	SSD-mobilnet-v2	0,81	0,80	0,46	0,83	0,02	0,76	0,81
Полнота	Efficient-D0	0,23	0,70	0,11	0,72	0,0001	0,89	0,57
	Efficient-D1	0,03	0,80	0,11	0,69	0,031	0,80	0,56
	Efficient-D2	0,22	0,68	0,04	0,75	0,066	0,73	0,57
	Faster-resnet50	0,57	0,77	0,85	0,76	0,039	0,70	0,79
	SSD-mobilnet-v2	0,21	0,69	0,44	0,75	0,012	0,79	0,61

При тестировании алгоритма кластеризации прогнозов обнаруженных объектов на первом этапе был осуществлен подбор минимального требуемого количества кадров. Для каждого кадра обнаруженные объекты записываются в массив предсказаний объектов возгорания, представленных *координатами углов* ограничивающей рамки и *вероятностью* принадлежности области, ограниченной рамкой, к искомому классу. В результате тестирования было определено количество кадров для анализа равное 4. Применение алгоритма на четырех кадрах повысило показания полноты на 3 %, при сохранившихся результатах по метрике точности. При этом реализация алгоритма позволила получить высокую эффективность объединения предсказаний обнаружения дыма, увеличив среднюю точность локализации найденных объектов на изображении на 6 % и составила по метрике IoU 62 % (рисунок 4).

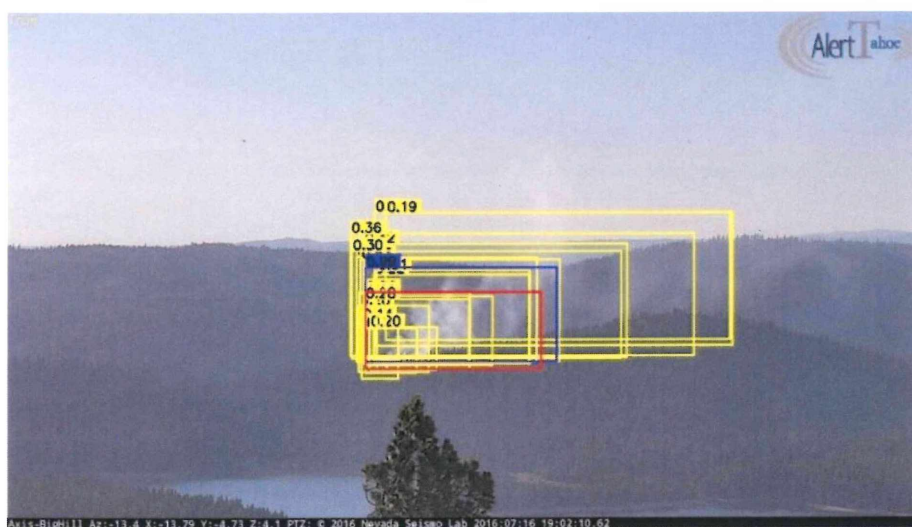


Рисунок 4 – Пример работы алгоритма кластеризации прогнозов обнаруженных объектов (боксы жёлтого цвета результаты предсказаний, бокс красного цвета – размеченный бокс, бокс синего цвета – работа алгоритма)

Блок 3 технологии обнаружения пожароопасных объектов в лесном массиве (рисунок 1) представлен двумя алгоритмами, а именно, алгоритмом подготовки данных и алгоритмом фильтрации обнаруженных областей задымления. Алгоритм подготовки данных по координатам, полученным из Блока 2 (рисунок 1), формирует последовательность фрагментов изображений из кадров исходной видеопоследовательности в количестве 4 штук. Полученная последовательность отправляется на анализ в алгоритм классификации обнаруженных объектов.

Для тестирования алгоритма фильтрации обнаруженных объектов были обучены 5 моделей сверточных нейронных сетей (рисунок 5). В результате тестирования (таблица 3) выбрана модель сверточной нейронной сети Inception-V3. Следующий этап заключался в оптимизации гиперпараметров гибридной архитектуры нейронной сети, состоящей из сверточной и рекуррентной нейронных сетей (таблица 4). Применение алгоритма фильтрации обнаруженных объектов увеличило показатель точности (Precision) на 12 % относительно результатов, полученных после обнаружения объекта. Точность системы по метрике классификационной точности (Accuracy) составила 95,6 %, показатель метрики F-score составил 94%.

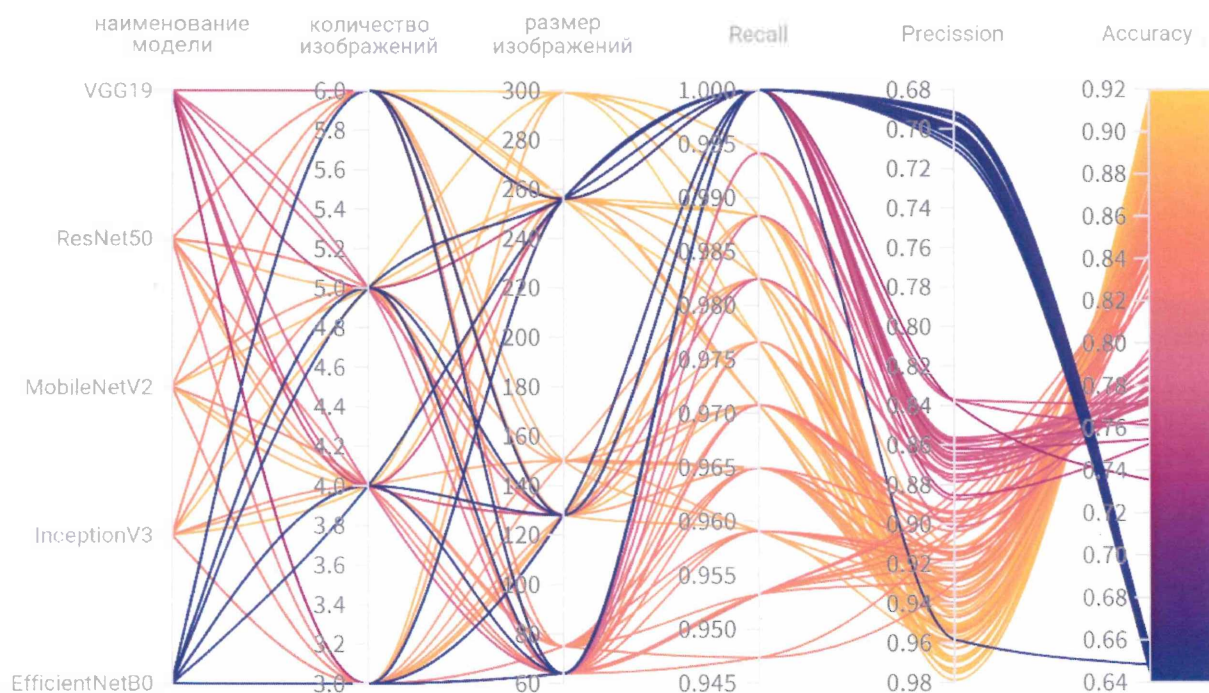


Рисунок 5 – График результатов обучения и подбор параметров моделей сверточных нейронных сетей

Таблица 3 – Результаты обучения моделей НС

Модель	Recall	Precision	Accuracy	F-score
VGG19	0,82	0,87	0,79	0,84
MobileNet-V2	0,94	0,92	0,91	0,93
ResNet50	0,94	0,91	0,90	0,92
Inception-V3	0,94	0,93	0,92	0,94
EfficientNet-B0	0,98	0,65	0,65	0,78

Таблица 4 – Параметры обучения гибридных нейросетей

Параметр сети	Характеристики
Размер входного изображения, px	64, 128, 256, 300
Коэффициент скорости обучения, у.е.	0,0001
Метод оптимизации	Adam
Тип функции активации	LeaklyReLU
Функция потерь	Бинарная кросс энтропия
Тип слоя РНС	GRU, LSTM
Размер слоя РНС, нейронов	128
Размер слоя FC, нейронов	128, 256, 512, 1024
Размер батча, шт.	12, 20, 24
Метод Dropout, у.е.	0,5

В целях оценки работоспособности и надежности предложенных алгоритмов проведено сравнение полученных результатов с аналогичными системами обнаружения, чьи характеристики представлены в открытом доступе. Результаты сравнения представлены в таблице 5, предложенная в диссертации технология обнаружения очага возгорания в лесном массиве демонстрирует более высокую точность обнаружения и меньшее количество ложных срабатываний.

Таблица 5 – Сравнение подходов обнаружения очага возгорания в лесном массиве

Подход		Точность, %	Полнота, %	Классификационная точность (Accuracy), %
1	Система обнаружения лесных пожаров Fuego	91,4	86,0	91,2
2	Система обнаружения лесных пожаров SmokeyNet	90,9	76,1	83,6
3	Система обнаружения лесных пожаров Multimodal SmokeyNet	87,0	71,0	80,0
4	Предложенная технология	95,4	97,2	95,6

Отметим, что применение разработанных алгоритмов позволяет в разных областях деятельности решить задачу выделения объектов с динамическими признаками в видеопоследовательности, адаптируя лишь параметры системы.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ВЫВОДЫ

1. Предложена технология анализа изображений в видеоряде позволяющей достичь высокого качества обнаружения объектов, на основе предложенных алгоритмов локализации и классификации с учетом статических и динамических данных. Предложенные в технологии алгоритмы могут быть применены с целью анализа изображений для широкого круга практических задач.
2. Реализован алгоритм выделения динамических признаков на последовательности кадров, основанный на алгоритмах цветовой обработки изображений и

операции вычитания фона. При реализации алгоритма осуществляется вычисление средней интенсивности пикселей кадров последовательности с последующим вычислением разности ключевого кадра с одним из исходных кадров. Использование разработанного алгоритма для модели обнаружения объектов Efficientdet-D0 позволило повысить значения метрики полноты (Recall) на 31 %, а метрики точности (Precision) на 21 %.

3. Реализован алгоритм, позволяющий объединять результаты нейросети в итоговое решение. Алгоритм выделяет плотные группы объектов данных, на основе оценки отношения площадей пересечения между соседними объектами, что отличает от других алгоритмов, базирующихся на фиксированном значении расстояния между объектами. Использование разработанного алгоритма позволяет повысить точность локализации на 6 % относительно альтернативных подходов.
4. Предложен алгоритм классификации обнаруженных объектов, состоящий из алгоритма подготовки последовательности изображений для анализа гибридной архитектурой нейронной сети. Использование гибридной архитектуры позволяет анализировать статические характерные для дыма признаки с учетом хронологии изменения объектов. Применение предложенного классификатора позволило увеличить показатель F-score на 9 %.
5. Результаты диссертационной работы были использованы при выполнении НИР грант УМНИК №16068ГУ/2020, грант РФФИ № 20-37-90055, грант ГЗ "Наука" №FFSWW-2020-0014. Внедрены в организациях в ООО "НТП КИБЕРЦЕНТР" (г. Томск), в ООО "Спотпаркинг" (г. Томск) и в учебном процессе в отделении информационных технологий Инженерной школы информационных технологий и робототехники ТПУ.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Публикации в изданиях, рекомендованных ВАК:

1. Laptev, N.V. Visualization System for Fire Detection in the Video Sequences // N.V. Laptev, V.V. Laptev, O.M. Gerget et al. // Scientific Visualization. – 2021. – Vol. 13. – No 2. – P. 1-9. (Scopus).
2. Лаптев, Н.В. Решение проблемы объединения прогнозов выделенных на изображении объектов дыма / Н.В. Лаптев, О.М. Гергет, А.А. Кравченко, В.В. Лаптев, Д.Ю. Колпашиков // Информационные и математические технологии в науке и управлении. – 2022. – № 4(28). – С. 136-143.
3. Лаптев, Н.В. Разработка библиотеки планирования и управления движениями коллаборативного робота Kuka iiwa / Н.В. Лаптев, А.А. Кравченко, О.М. Гергет // Информационные и математические технологии в науке и управлении. – 2022. – № 2(26). – С. 161-168.
4. Лаптев, Н.В. Автоматическое управление непрерывным роботом с использованием алгоритма FABRIK / В.В. Данилов, Д.Ю. Колпашиков, Н.В. Лаптев // Моделирование, оптимизация и информационные технологии. – 2019. – Т. 7. – № 4(27). – С. 1-2.

5. Laptev, N.V. Use of semi-synthetic data for catheter segmentation improvement // V.V. Danilov, D.Y. Kolpashchikov, N.V. Laptev, O.M. Gerget et al. // Computerized Medical Imaging and Graphics. – 2023. – Vol. 106. – P. 102188. (Scopus).

Свидетельства о регистрации интеллектуальной деятельности:

1. Патент на изобретение № 2721769. Стенд для контроля контурных перемещений гибкого манипулятора/ Лаптев Н.В., Колпашиков Д.Ю., Гергет О.М. и др. // Оpubл. 22.05.2020, Бюл. № 15. Заявка № 2019127060 от 28.08.2019 г. Решение о выдаче патента от 28.08.2019 г.
2. Свидетельство о государственной регистрации программы для ЭВМ № 2022611227 (RU); заявка № 2022610413 от 13.01.2022, дата рег. 21.01.2022 // Лаптев Н.В., Лаптев В.В., Гергет О.М. Обнаружение пожароопасных объектов в лесном массиве на основе динамических признаков.
3. Свидетельство о государственной регистрации программы для ЭВМ № 2022684945 (RU); заявка № 2022685130 от 20.12.2022, дата рег. 20.12.2022 // Лаптев Н.В., Гергет О.М., Колпашиков Д.Ю., Кравченко А.А., Лаптев В.В. Объединение прогнозов обнаруженных областей задымления.
4. Свидетельство о государственной регистрации программы для ЭВМ № 2022619805 (RU); заявка № 2022618392 от 05.05.2022, дата рег. 26.05.2022 // Сонькин Д.М., Плешков А.Г., Лаптев Н.В. и др. Программа для ЭВМ "Многофункциональный мобильный комплекс мониторинга, ситуационного анализа и оповещения труднодоступных объектов с использованием беспилотных летательных аппаратов и ГИС-технологий (ММК МАО)".

Публикации в других изданиях:

1. Laptev, N. Integrating traditional machine learning and neural networks for image processing / N. Laptev, V. Laptev, O. Gerget, D. Kolpashchikov // CEUR Workshop Proceedings: 31, Nizhny Novgorod, 27–30 сентября 2021 года. – Nizhny Novgorod, 2021. – P. 896-904. (Scopus).
2. Лаптев, Н.В. Обнаружение пожароопасных участков на кадрах видеоряда посредством классификации / Н.В. Лаптев, В.В. Лаптев // Инновационные технологии: теория, инструменты, практика. – 2020. – Т. 1. – С. 62-68.
3. Laptev, N.V. Visual shape and position sensing algorithm for a continuum robot / R.A. Manakov, D.Y. Kolpashchikov, N.V. Laptev et al. // IOP Conference Series: Materials Science and Engineering: 14, Tomsk, 14–17 октября 2019 года. – Tomsk, 2021. – P. 012066. (Scopus).
4. Laptev, N.V. Boosting segmentation accuracy of the deep learning models based on the synthetic data generation / V.V. Danilov, O.M. Gerget, N.V. Laptev et al. // International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives: 4, Moscow, 26–28 апреля 2021 года. – Moscow, 2021. – P. 33-40. (Scopus).