Обнаружение и классификация пожаров с помощью нейросетей

19 октября 2024 г.

**Аннотация**

Обнаружение и мониторинг пожаров являются одними из важнейших задач в современных системах безопасности и защиты окружающей среды. Ежегодно лесные пожары наносят огромный урон экосистемам, экономике и человеческой жизни. В связи с этим развитие технологий раннего обнаружения пожаров становится приоритетной задачей для многих стран, особенно с учетом роста масштабов и частоты природных катастроф.

Вовремя обнаруженный пожар позволяет минимизировать ущерб и предотвратить человеческие жертвы. Современные технологии, такие как дроны, оснащённые тепловизорами и обычными камерами, могут предоставить богатые данные для анализа. Однако сложность обработки этих данных заключается в необходимости выявления признаков пожара в условиях переменных освещённости, различной погоды и изменяющихся текстур объектов. Для решения этой задачи активно используются методы искусственного интеллекта, в частности нейронные сети.

Цель данной работы — рассмотреть основные методы обнаружения пожаров с использованием данных, полученных с дронов и тепловизоров, а также проанализировать существующие архитектуры нейросетей, применяемые для решения данной задачи.

**Введение**

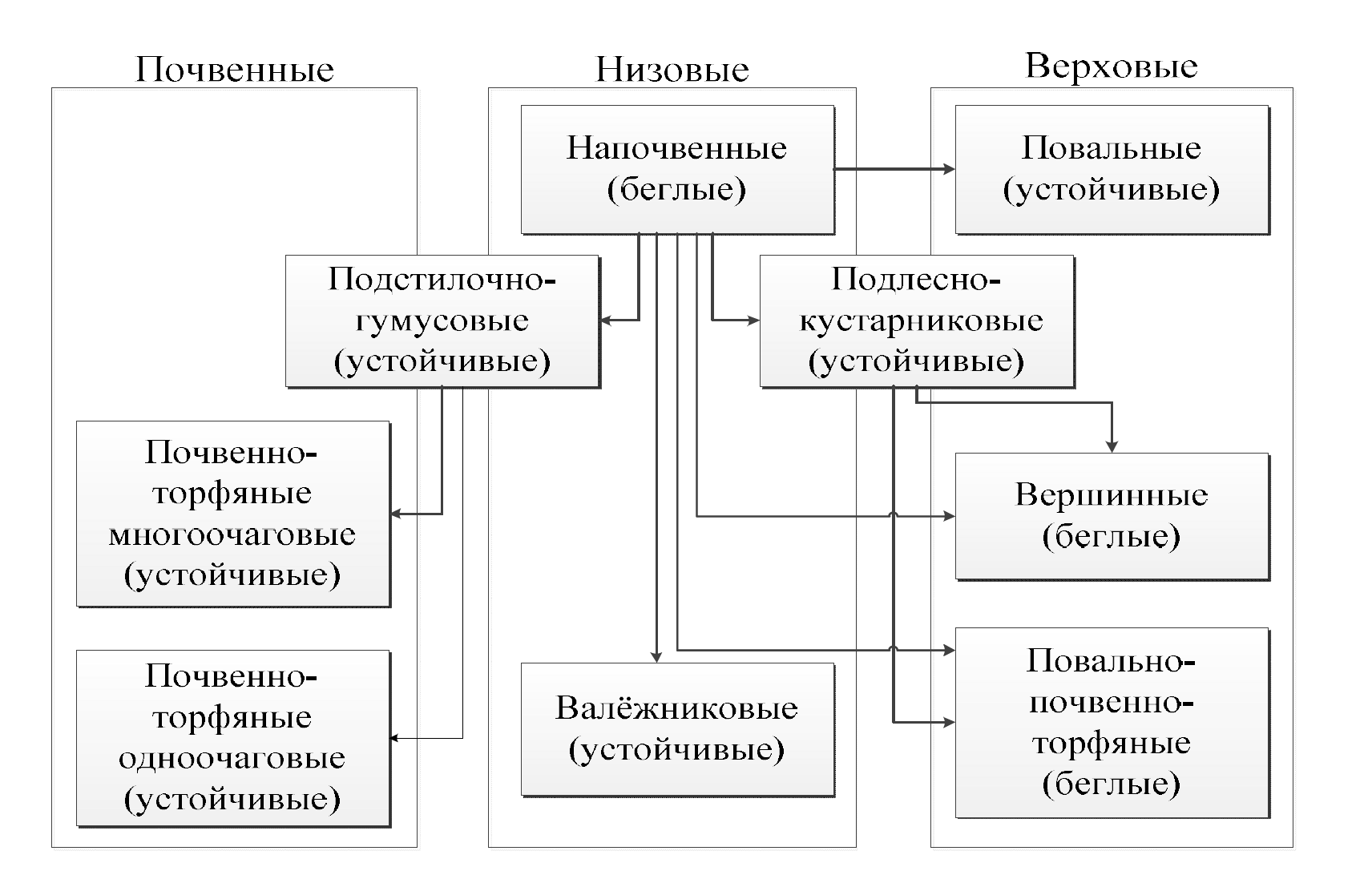
Классификация пожаров играет важную роль в нейросетевом обнаружении, так как она позволяет улучшить точность детекции, адаптировать модели под различные условия, и более эффективно использовать данные с дронов. Это обеспечивает не только раннее обнаружение пожаров, но и позволяет принимать более обоснованные решения по их предотвращению и тушению.

Рис. 1 - Классификация лесных пожаров Н.П. Курбатского (1964)

В классификации пожаров, предложенной Курбатским в 1964 году, пожары делятся на несколько типов, каждый из которых сопровождается разными визуальными признаками, такими как дым или огонь.

**Почвенные:**

* **Признаки:** Эти пожары распространяются по лесной подстилке — сухим листьям, хвое, веткам и другим органическим материалам на поверхности земли.
* **Визуальные проявления:** Обычно сопровождаются плотным дымом, но огонь часто не виден с воздуха, особенно на ранних стадиях. Этот тип пожаров может тлеть на небольшой глубине, выделяя значительные объемы дыма, но без явных языков пламени.

**Наземные (низовые) пожары:**

* **Признаки**: Пожары, которые распространяются по низкорослой растительности — кустарникам, траве.
* **Визуальные проявления**: Часто сопровождаются как дымом, так и открытым огнём, особенно если ветер усиливает пламя. Огонь может быть заметен, но основное внимание привлекает дым, особенно на больших расстояниях.

**Верховые пожары:**

* **Признаки**: Распространяются по кронам деревьев, что делает их самыми интенсивными и опасными пожарами.
* **Визуальные проявления**: Кроновые пожары характеризуются сильным огнём, который распространяется через верхушки деревьев. Визуально их легко распознать по ярким языкам пламени, которые видны даже с большого расстояния, а также по черному густому дыму, поднимающемуся высоко в атмосферу.

Таким образом, почвенные пожары в основном сопровождаются сверху дымом, а верховые и наземные пожары могут сопровождаться как дымом, так и открытым огнем.

Существует множество различных датасетов и математических методов, которые применяются для нейросетевого обнаружения пожаров. Эти датасеты содержат изображения пожаров, полученные с различных источников, включая дроны, спутники и тепловизоры, что позволяет моделям нейронных сетей обучаться на разнообразных данных. Кроме того, разработано множество математических методов обработки изображений, таких как сегментация, фильтрация, а также сложные алгоритмы глубокого обучения, которые помогают повысить точность и надежность обнаружения пожаров в различных условиях.

**Набор данных**

В данной работе использовался набор данных DFireDataset.

DFireDataset — это набор изображений, разработанный для обнаружения пожаров и дыма с использованием методов машинного обучения. Датасет содержит более 21 000 изображений, на которых зафиксированы как пожары, так и дым, и классифицирует их по четырем категориям:

1. Только огонь (1,164 изображения)
2. Только дым (5,867 изображений)
3. Огонь и дым (4,658 изображений)
4. Без пожара и дыма (9,838 изображений)

В дополнение к этому, набор данных включает **bounding boxes** (bounding box — это прямоугольная область, которая используется для выделения объектов на изображениях в задачах компьютерного зрения) для каждого объекта, обозначающие местоположение огня и дыма, что делает его полезным для обучения алгоритмов детекции объектов. Все изображения аннотированы в формате YOLO, с координатами, нормализованными от 0 до 1. Однако также предоставлена функция для конвертации координат в пиксели, если это необходимо. DFireDataset используется для задач детекции пожаров с применением глубоких нейронных сетей.

А B

Рис. 2 -Примеры объектов набора данных; А - пример 1, B - пример 2

****

Рис. 3 - Изображения с bounding boxes

**Выбор методов для обнаружения пожаров**

| Критерий | CNN(сверточные нейросети) | Классические методы машинного обучения | Другие нейросетевые архитектуры(например, полносвязные сети) |
| --- | --- | --- | --- |
| **Автоматическое извлечение признаков** | CNN автоматически выделяют признаки на разных уровнях изображения, что упрощает обучение и делает метод более универсальным. | Требуют ручного выделения признаков, что делает процесс более трудоёмким и менее гибким. | Полносвязные сети или RNN неэффективны для обработки изображений из-за своей архитектуры; требуют большой объём вычислений. |
| **Точность и масштабируе-мость** | CNN демонстрируют высокую точность в обнаружении объектов и масштабируются для различных типов данных. | Точность зависит от качества выделенных признаков, что делает их менее эффективными при сложных данных. | Полносвязные сети плохо масштабируются для задач обработки изображений; менее точны в детекции. |
| **Работа с большими объёмами данных** | CNN эффективно обучаются на больших датасетах, что улучшает их способность выявлять сложные паттерны и объекты. | Могут страдать от проблем с переобучением на больших данных или сложных изображениях; требуют уменьшения размерности данных. | Другие нейросети плохо справляются с большими изображениями и высокоразмерными данными, требуют предобработки. |
| **Устойчивость к шуму и различным условиям** | CNN устойчивы к различным условиям съёмки, благодаря множественным свёрточным слоям и уровням абстракции признаков. | Менее устойчивы к изменяющимся условиям, так как выделенные вручную признаки часто привязаны к конкретным условиям освещения или текстур. | Полносвязные сети и RNN менее устойчивы к изменяющимся условиям, так как не могут эффективно работать с изображениями. |
| **Интеграция с тепловизионными данными** | CNN могут легко интегрировать мультиспектральные данные, такие как RGB и тепловизионные изображения, что помогает улучшить результаты обнаружения пожаров. | Классические методы требуют сложных процессов предобработки и выделения признаков для работы с мультиспектральными данными. | Другие нейросети неэффективны в интеграции таких данных без значительных изменений архитектуры. |
| **Производительность в реальном времени** | Одноступенчатые CNN-детекторы, такие как YOLO, обеспечивают высокую скорость работы и точность, что делает их идеальными для применения в реальном времени, например, при мониторинге с дронов. | Классические методы обычно медленнее и могут не подходить для задач в реальном времени, особенно при больших объёмах данных. | Полносвязные сети или RNN значительно медленнее и не подходят для задач в реальном времени, связанных с обработкой изображений. |
| **Гибкость и адаптируемость** | CNN легко адаптируются к новым типам данных, типам пожаров и различным источникам. | Классические методы менее гибкие, так как требуют ручной настройки параметров и выделения признаков для новых данных. | Полносвязные сети менее гибкие и требуют значительных изменений при адаптации к новым данным. |

Табл. 1 - Таблица сравнения методов распознавания пожаров

**Описание Табл. 1:**

**1. Автоматическое извлечение признаков**

* **CNN (Свёрточные нейронные сети)**:  
  CNN автоматически выделяют признаки на различных уровнях изображения, начиная от простых контуров до сложных объектов. Это одно из ключевых преимуществ CNN, поскольку автоматизация извлечения признаков снижает необходимость в ручной настройке и повышает адаптивность модели к новым типам данных. Свёртки и пулинг позволяют модели «увидеть» различия на разных уровнях абстракции изображения, что делает её более точной и гибкой.
* **Классические методы машинного обучения**:  
  В классических методах, таких как SVM или случайные леса, используется ручное выделение признаков. Методы, такие как HOG (гистограмма направленных градиентов) и LBP (локальные бинарные шаблоны), требуют настройки параметров и выбора подходящих признаков. Это делает их менее гибкими, так как ручной выбор признаков может не учесть сложные взаимодействия и корреляции в данных.
* **Другие нейросетевые подходы (полносвязные сети, RNN)**:  
  Полносвязные нейросети (MLP) не предназначены для автоматического извлечения признаков и работают только с теми признаками, которые были заранее выделены вручную. Это делает их менее эффективными для обработки изображений, так как они не используют пространственную структуру данных. RNN же лучше подходят для временных последовательностей, а не для работы с изображениями.

**2. Точность и масштабируемость**

* **CNN**:  
  CNN демонстрируют высокую точность в задачах обнаружения объектов, поскольку они могут выявлять сложные взаимосвязи между пикселями. Эти модели также легко масштабируются для различных типов данных: RGB, тепловизионные изображения или комбинированные мультиспектральные данные. Использование многослойных свёрточных слоёв позволяет CNN адаптироваться к различным уровням детализации данных.
* **Классические методы машинного обучения**:  
  Точность работы классических методов зависит от качества ручного выделения признаков. Чем лучше выделены признаки, тем выше будет точность модели. Однако масштабируемость таких методов ограничена, так как при увеличении количества данных или появлении новых типов данных (например, тепловизионные снимки) требуется пересмотр и перенастройка алгоритмов.
* **Другие нейросетевые подходы**:  
  Полносвязные сети плохо масштабируются для задач с изображениями, так как их структура не позволяет эффективно обрабатывать большие пространственные данные. RNN также не предназначены для обработки изображений, их точность в подобных задачах низка.

**3. Работа с большими объёмами данных**

* **CNN**:  
  CNN эффективно обрабатывают большие датасеты, поскольку свёрточные слои позволяют выявлять сложные паттерны в данных. При этом количество параметров не растёт экспоненциально с увеличением разрешения изображения, что делает их применимыми для больших наборов данных.
* **Классические методы машинного обучения**:  
  Классические методы машинного обучения склонны к переобучению при работе с большими объёмами данных, особенно если данные имеют высокую размерность. Для снижения этого эффекта часто приходится уменьшать размерность или выбирать только ключевые признаки, что может привести к потере полезной информации.
* **Другие нейросетевые подходы**:  
  Полносвязные сети не приспособлены для работы с большими изображениями, так как количество параметров резко возрастает с увеличением количества пикселей, что делает такие модели неэффективными. RNN же лучше справляются с временными последовательностями данных, но не с изображениями.

**4. Устойчивость к шуму и различным условиям**

* **CNN**:  
  CNN благодаря многослойной архитектуре устойчивы к изменениям внешних условий (например, изменению освещённости, наличию шума или задымления). Уровни абстракции позволяют сети игнорировать незначительные изменения и сосредоточиться на ключевых признаках, что делает их надёжными в условиях, где другие методы могут потерпеть неудачу.
* **Классические методы машинного обучения**:  
  Классические методы, такие как HOG и LBP, менее устойчивы к шуму и изменению условий. Если признаки выделены вручную и они зависят от конкретных условий (например, освещённости), то модель может работать хуже, если условия меняются (например, присутствие дыма или смена времени суток).
* **Другие нейросетевые подходы**:  
  Полносвязные сети и RNN также плохо адаптируются к изменениям условий, поскольку они не могут автоматически выделять признаки и требуют от разработчика заранее подготовить данные, что ограничивает их гибкость и устойчивость.

**5. Интеграция с тепловизионными данными**

* **CNN**:  
  CNN легко интегрируют мультиспектральные данные (например, изображения в видимом и инфракрасном диапазоне). Это позволяет использовать несколько каналов данных для более точного обнаружения пожаров, особенно в условиях плохой видимости (например, ночью или при задымлённости).
* **Классические методы машинного обучения**:  
  Для работы с мультиспектральными данными классические методы требуют сложной предобработки и настройки для каждого типа данных. Это делает их менее гибкими в задачах, где важно комбинировать различные источники информации (например, видимые и тепловые изображения).
* **Другие нейросетевые подходы**:  
  Полносвязные сети и RNN не предназначены для эффективной интеграции данных разных типов. Для работы с такими данными потребуется значительная модификация их архитектуры.

**6. Производительность в реальном времени**

* **CNN**:  
  Одноступенчатые детекторы, такие как YOLO, на основе CNN обеспечивают высокую скорость работы при сохранении высокой точности. Это делает их отличным выбором для задач реального времени, таких как мониторинг пожаров с дронов, где необходимо быстро обрабатывать входящие данные.
* **Классические методы машинного обучения**:  
  Классические методы, такие как SVM или случайные леса, обычно медленнее работают в задачах реального времени. Это связано с тем, что этапы выделения признаков и классификации могут быть вычислительно затратными.
* **Другие нейросетевые подходы**:  
  Полносвязные сети и RNN значительно медленнее для задач реального времени, особенно если речь идёт о больших изображениях. Их архитектуры не приспособлены для быстрого анализа данных в режиме реального времени.

**7. Гибкость и адаптируемость**

* **CNN**:  
  CNN легко адаптируются к новым данным и различным типам пожаров. Модель можно быстро переобучить или дообучить для работы с новыми условиями (например, другим типом ландшафта или изменившимися внешними факторами).
* **Классические методы машинного обучения**:  
  Классические методы менее гибкие, так как для каждого нового типа данных или условий необходимо пересматривать подходы к выделению признаков и их обработке. Это делает их менее эффективными в динамично меняющихся условиях.
* **Другие нейросетевые подходы**:  
  Полносвязные сети и RNN также менее гибкие, так как для их адаптации к новым данным требуется значительная модификация модели.

Таким образом, свёрточные нейронные сети (CNN) представляют собой наиболее эффективное решение для обнаружения пожаров среди классических методов и других типов нейросетей.

**Выбор сверточной нейросети**

В данной работе для обнаружения пожаров была выбрана нейросеть YOLO v.8.

Далее показано сравнение данной нейросети с актуальными конкурентами.

**YOLOv8** (You Only Look Once, версия 8) была выбрана для задачи обнаружения пожаров по нескольким ключевым причинам, которые делают её более подходящей по сравнению с другими нейросетевыми подходами и моделями на основе свёрточных нейронных сетей (CNN).

**Одноступенчатая архитектура и высокая скорость**

YOLOv8 представляет собой одноступенчатый детектор объектов, что означает, что она одновременно выполняет детекцию и классификацию объектов, избегая разделения на несколько этапов, как это происходит в двухступенчатых моделях. Это позволяет значительно ускорить обработку изображений, что особенно важно для задач, требующих работы в реальном времени, таких как мониторинг пожаров с использованием дронов. В отличие от моделей вроде Faster R-CNN или Mask R-CNN, которые требуют больше времени для анализа каждого кадра из-за сложных двухступенчатых процессов, YOLOv8 может выполнять детекцию объектов практически мгновенно.

**Точность обнаружения объектов**

YOLOv8 значительно улучшила точность по сравнению с предыдущими версиями YOLO, а также с другими аналогичными моделями. Она обеспечивает высокую точность в обнаружении и локализации объектов, таких как огонь и дым, даже в условиях недостаточного освещения или сильного задымления. Эта точность играет ключевую роль в задаче обнаружения пожаров, где важно быстро и точно определить местоположение очага возгорания для его дальнейшей ликвидации.

**Масштабируемость и гибкость**

YOLOv8 легко адаптируется к различным типам данных и может работать с мультиспектральными изображениями, такими как фото с обычных камер и тепловизоров. Это важно в условиях мониторинга пожаров, где данные могут поступать с разных источников, и нужно уметь их объединять для создания более полной картины. Возможность эффективной работы с различными типами данных делает YOLOv8 универсальной моделью для решения задач, связанных с обнаружением пожаров.

**Компактность и эффективность модели**

YOLOv8 отличается меньшими размерами и большей эффективностью по сравнению с другими сложными моделями на основе CNN. Это делает её пригодной для использования в мобильных и встроенных системах, что особенно важно для дронов, где вычислительные ресурсы ограничены. Хотя такие модели, как EfficientDet или RetinaNet, демонстрируют хорошие результаты, они могут оказаться менее эффективными для задач, требующих быстрой обработки данных и минимизации ресурсоёмкости, что критично при использовании на дронах.

**Устойчивость к шуму и изменяющимся условиям**

YOLOv8 обладает высокой устойчивостью к различным условиям съёмки, включая низкую освещённость и задымление. По сравнению с предыдущими версиями YOLO и многими другими моделями, YOLOv8 показывает лучшие результаты в условиях шума и изменяющихся факторов окружающей среды, что делает её более надёжной для обнаружения пожаров в реальных условиях.

**Поддержка работы в реальном времени**

Одним из ключевых преимуществ YOLOv8 является её высокая производительность в задачах реального времени. По сравнению с двухступенчатыми моделями, такими как Faster R-CNN, YOLOv8 выполняет обработку изображений значительно быстрее, что позволяет оперативно обнаруживать пожары. Это особенно важно при работе с видеопотоками с дронов, где необходимо быстро реагировать на угрозы.

**Сравнение с другими актуальными нейросетями для обнаружения пожаров**

1. **Faster R-CNN (2015)** Faster R-CNN представляет собой двухступенчатый детектор объектов с высокой точностью, однако его сложная архитектура и большая вычислительная сложность делают его менее подходящим для задач реального времени. Faster R-CNN обрабатывает кадры значительно медленнее, чем YOLOv8, что снижает его применимость для задач, требующих быстрого реагирования.
2. **EfficientDet (2020)** EfficientDet — это современная модель, которая балансирует между точностью и вычислительной эффективностью. Однако она требует больше вычислительных ресурсов и сложнее в настройке для работы в реальном времени, по сравнению с YOLOv8. EfficientDet хороша для задач, требующих максимальной точности, но уступает YOLOv8 в условиях ограниченных ресурсов, таких как мониторинг пожаров с дронов.
3. **RetinaNet (2017)** RetinaNet известна своим подходом к решению проблемы дисбаланса между положительными и отрицательными примерами, что позволяет ей хорошо детектировать малые объекты, такие как небольшие очаги огня. Однако она использует двухступенчатую архитектуру, что делает её медленнее, чем YOLOv8. RetinaNet может быть эффективной для точной локализации, но её скорость недостаточна для задач реального времени.
4. **DETR (DEtection TRansformer, 2020)** DETR — это новая архитектура на основе трансформеров, показывающая хорошие результаты в задачах детекции объектов. Тем не менее, она медленнее YOLOv8 и требует значительно больше вычислительных ресурсов для тренировки. DETR интересен для будущих исследований, однако для текущих задач реального времени он менее пригоден.
5. **SSD (Single Shot Detector, 2016)** SSD, как и YOLO, является одноступенчатым детектором, способным обрабатывать данные в реальном времени. Однако YOLOv8 демонстрирует более высокую точность при сопоставимой скорости. SSD остаётся популярным, но в современных приложениях предпочтение всё чаще отдаётся новым версиям YOLO из-за их улучшенных показателей точности и гибкости.

**Заключение**

В данной работе были рассмотрены особенности классификации пожаров и наиболее популярные методы их обнаружения. Кроме того, было проведено сравнение методов решения поставленной задачи, и была выбрана сверточная нейросеть YOLO v8.

Выбор YOLOv8 для задачи обнаружения пожаров основывается на её сбалансированности между скоростью и точностью, что критично в задачах мониторинга в реальном времени. Хотя другие современные модели, такие как EfficientDet, RetinaNet и Faster R-CNN, могут показывать высокие результаты в определённых условиях, YOLOv8 выигрывает за счёт своей скорости, компактности и способности адаптироваться к различным источникам данных, что делает её лучшим решением для мониторинга пожаров с дронов.

**Ссылки**

1. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). "You Only Look Once: Unified, Real-Time Object Detection." <https://arxiv.org/abs/1506.02640>
2. Ren, S., He, K., Girshick, R., & Sun, J. (2015). "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks." <https://arxiv.org/abs/1506.01497>
3. Tan, M., Pang, R., & Le, Q. V. (2020). "EfficientDet: Scalable and Efficient Object Detection." <https://arxiv.org/abs/1911.09070>
4. Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollar, P. (2017). "Focal Loss for Dense Object Detection." <https://arxiv.org/abs/1708.02002>
5. Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020). "End-to-End Object Detection with Transformers." <https://arxiv.org/abs/2005.12872>
6. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). "SSD: Single Shot Multibox Detector." <https://arxiv.org/abs/1512.02325>
7. LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). "Gradient-Based Learning Applied to Document Recognition." Proceedings of the IEEE, 86(11), 2278-2324. <https://www.researchgate.net/publication/2985446_Gradient-Based_Learning_Applied_to_Document_Recognition>
8. Fan, Linwei & Zhang, Fan & Fan, Hui & Zhang, Caiming. (2019). Brief review of image denoising techniques. Visual Computing for Industry, Biomedicine, and Art. 2. 10.1186/s42492-019-0016-7.
9. DFireDataset: <https://github.com/gaiasd/DFireDataset>
10. Залесов А.С. «Классификация лесных пожаров». https://elar.usfeu.ru/bitstream/123456789/248/3/Zalesov\_A.S.\_Klassifikaciya\_lesnix\_pozgarov.pdf