

Обнаружение пожаров

26 октября 2024 г.

Наборы данных

Набор данных FLAME2: Разработан лабораторией AI-SENDS Клемсонского университета, включает синхронизированные изображения RGB и инфракрасного (IR) диапазонов, снятые в процессе контролируемых поджогов. Набор данных насчитывает более 53,000 помеченных пар кадров RGB и IR, размеченных по категориям «Огонь/Нет огня» и «Дым/Нет дыма». Также включены дополнительные данные, такие как погодные условия и 3D-карты, что улучшает возможности для разработки и тестирования моделей глубокого обучения.

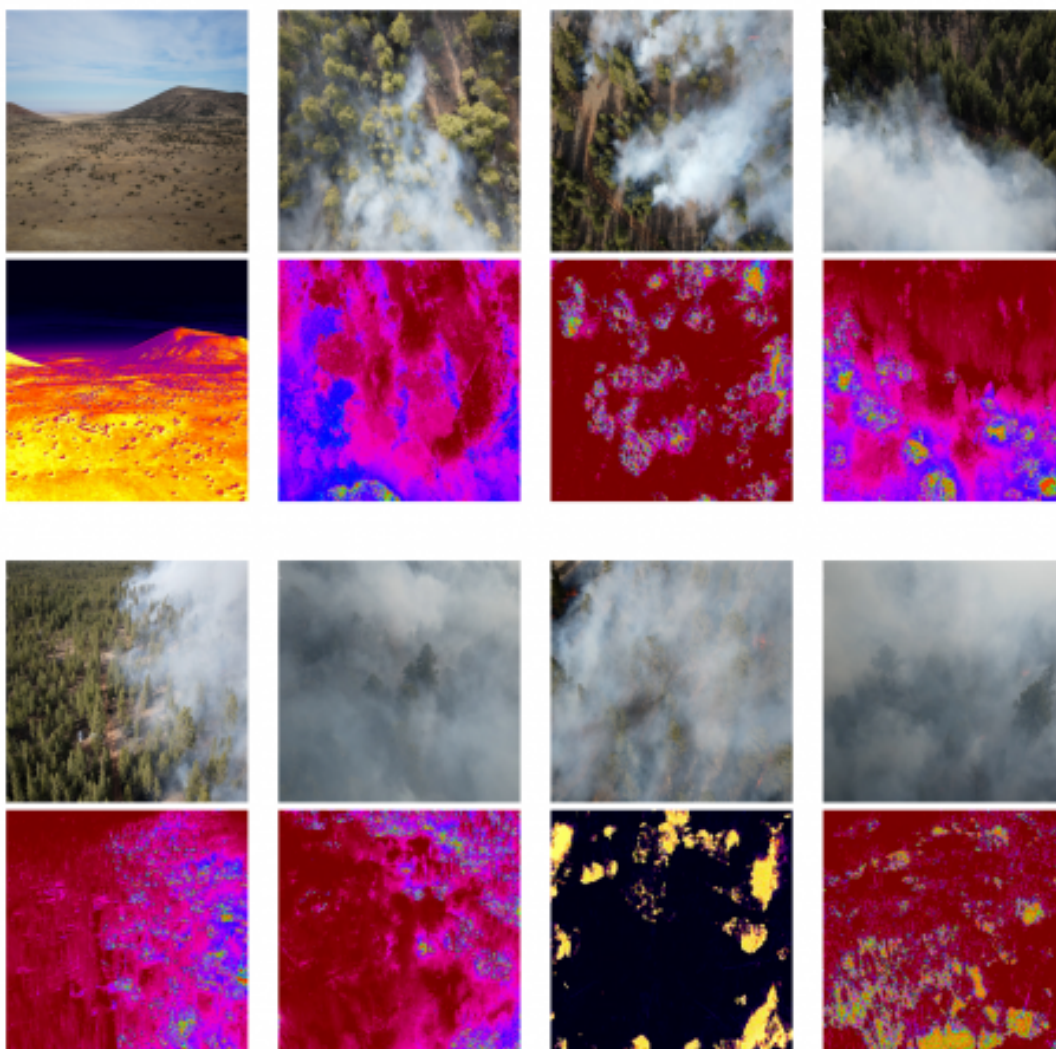


Рис. 1 - Пример объектов в наборе Flame 2

Создание собственного датасета для распознавания белого дыма/тумана/облаков

Вместо использования единого готового набора данных, можно создать свой датасет, объединив изображения из нескольких наборов данных и дополнительно собирая метеорологические данные для улучшения фильтрации. Так как не удалось найти датасет, хорошо подходящий для распознавания одновременно дыма, тумана и облаков, предлагается следующее решение:

1. Сбор данных для каждого класса (дым, туман, облака):

Дым: используем изображения из FIRESENSE, MODIS Fire and Smoke Dataset или FLAME2 для получения тепловых и RGB-снимков.

Туман и облака:

1 подход :

Добавим данные из DWIC (можно собрать данные на обычные и тепловизионные камеры в условиях высокой влажности). Если нужные тепловизионные данные отсутствуют в одном или нескольких датасетах, можно использовать генеративные методы, такие как CycleGAN или Pix2Pix, для синтеза тепловых изображений из видимых. Обучим CycleGAN или Pix2Pix на этом датасете для генерации тепловизионных данных из RGB изображений. Этот подход позволяет применять обученную сеть к любому RGB изображению, получая тепловой аналог. Эта модель не рассчитает точную температуру, но сможет научиться передавать особенности тепловых карт на основе большого количества примеров тепловых и обычных изображений.

2 подход :

Для создания датасета, где каждому изображению с видимым спектром соответствует тепловизионное изображение, можно организовать процесс сбора данных и аннотации вручную. Для каждой пары «видимое изображение — тепловизионное изображение» нужно убедиться, что объекты точно совпадают.

3 подход :

Симуляция тепловых изображений в моделировании. Если мы работаем с ограниченными сценами (например, с моделями тумана, дыма и облаков в определенных условиях), можно использовать **3D-симуляторы** (например, Unity

или Unreal Engine с шейдерами) для создания тепловых карт.

- Настройка сцены, подобной реальной обстановке, используем условные температурные значения.
- Применим к сцене шейдер, который перекрашивает объекты в зависимости от температуры, создавая тепловую карту.
- Используем эти сгенерированные данные как псевдотепловые изображения для тренировки модели.

Выбранный подход: создание датасета с ручной аннотацией пар видимых и тепловизионных изображений, так как он обеспечивает наиболее точное соответствие данных для обучения модели.

Создание вручную аннотированных пар изображений позволяет гарантировать, что видимые и тепловизионные изображения одного объекта идеально совпадают. Это особенно важно для задач, где малейшие смещения могут привести к ошибкам классификации.

Такой подход минимизирует влияние искажений, вызванных несовершенством синтетических методов или симуляций, например, использованием CycleGAN или 3D-моделирования.

Но, если объем данных ограничен и есть значительные временные ограничения, можно дополнить данные генеративными методами (подход 1) или симуляциями (подход 3). Однако они должны использоваться как вспомогательные, а не основные источники данных.

2. Аугментация и аннотация данных:

Создадим дополнительные изображения через аугментацию: измените яркость, контрастность, добавьте шум и атмосферные искажения, чтобы имитировать реальные условия.

Аннотируем объекты по классам ("дым", "туман", "облака") с указанием соответствующих координат и добавим метки для тепловых и видимых изображений для удобной синхронизации в обучении.

3. Добавление метеоданных:

Если есть возможность, собираем метеоданные параллельно с изображениями.

Например, можно использовать открытые метеоисточники (API, такие как OpenWeather или локальные погодные датчики), чтобы привязать данные о температуре и влажности к каждому снимку.

Подходы к использованию RGB и тепловых изображений в YOLO

Чтобы YOLO v8 могла эффективно работать с данными, включающими обычные и тепловые изображения, требуется немного изменить подход к подготовке и обучению модели. YOLO обычно обрабатывает RGB-изображения, поэтому для работы с дополнительным тепловым каналом можно использовать несколько техник.

1. Совмещение RGB и тепловых данных в один канал:

Можно конкатенировать RGB и тепловое изображение, создавая, по сути, "четырёхканальное" изображение. Например, третий канал RGB может быть заменен на тепловой. Однако YOLO v8 изначально не поддерживает такие данные, и потребуются небольшая модификация архитектуры модели для обработки теплового канала наряду с RGB.

При подготовке выборки каждое изображение должно содержать и RGB, и тепловой канал. Обычно тепловые данные нормализуются, чтобы значения были сопоставимы с диапазоном RGB (0–255).

YOLO v8, как правило, обучается на трехканальных изображениях, поэтому потребуется изменить входной слой модели для работы с четырьмя каналами. Это может потребовать обновления настроек свертки, чтобы поддерживать корректный расчет параметров на каждом этапе.

Совмещенное представление может улучшить способность модели распознавать объекты, полагаясь на оба источника информации, однако требует более тщательной настройки, так как увеличенный объем данных повышает вычислительные требования.

2. Отдельная обработка RGB и тепловых изображений с последующим объединением признаков:

В этом случае две версии изображения (RGB и тепловая) подаются в модель отдельно, и их признаки объединяются на определенном уровне свертки. Это требует либо создания кастомного слоя для объединения признаков, либо использования ансамблевой модели, где каждый подмодель обучается на своем типе данных, а затем совмещаются их результаты.

Архитектура YOLO v8 может быть модифицирована так, чтобы каждая ветвь модели обрабатывала свой тип данных независимо. Для этого создаются два параллельных потока (отдельные сети) — один для RGB, второй для теплового изображения. На определенном этапе (например, после нескольких сверточных слоев) эти потоки объединяются.

Признаки RGB и теплового изображения можно объединить с помощью методов конкатенации или сложения слоев. Например, после объединения они могут быть обработаны несколькими свертками, чтобы создать конечное представление объекта.

Такой подход позволяет каждому типу данных быть независимым, что снижает вероятность "подавления" теплового сигнала из-за более сильного сигнала RGB, и может повысить точность на многомодальных изображениях (RGB + тепловое).

3. Использование теплового изображения как дополнительной аугментации:

В некоторых случаях тепловое изображение можно рассматривать как дополнительный источник информации для аугментации, что повышает устойчивость модели. Например, можно включить тепловую карту в качестве дополнительного фона или наложения на обычное изображение.

Можно накладывать тепловую карту на RGB-изображение в виде полупрозрачного слоя. Это позволит модели получать некоторую информацию о температуре, даже если изображение остается в стандартном формате RGB.

Преимущества и недостатки: Этот подход не требует изменения архитектуры модели и может легко интегрироваться в стандартный процесс обучения. Однако он менее точен, так как тепловая информация подается модели не напрямую, а через визуальные эффекты.

4. Построение составного датасета:

Датасет может быть организован так, чтобы пара "RGB-изображение + тепловое изображение" использовалась в одной выборке, где каждое изображение хранит свою метку. Затем YOLO обучается на этом составном наборе данных, что позволяет модели адаптироваться к разным типам входных данных.

Выбранный вариант: Отдельная обработка RGB и тепловых изображений с последующим объединением признаков.

Пояснение:

Этот подход позволяет сохранить независимость обработки каждого типа данных, снижая вероятность "подавления" тепловой информации RGB-данными. Архитектура с двумя параллельными потоками улучшает точность распознавания за счет использования многомодальной информации. Она также лучше адаптируется к задачам, где требуется учитывать разницу в визуальных характеристиках тепловых и RGB-изображений.

Рассчет объема выборки

Для обучения модели YOLO для задачи распознавания дыма, тумана и облаков можно определить подходящий размер выборки, опираясь на следующие аспекты:

1. **Рекомендации по размеру выборки:** Специалисты по машинному обучению часто указывают, что модели для детектирования объектов (например, YOLO) требуют большого объема данных, особенно при работе с похожими объектами, такими как дым, туман и облака. Точных правил для объема данных нет, но оптимальным считается набор данных с тысячами размеченных изображений. Например, источники указывают, что для достижения хороших результатов требуется как минимум 5 000–10 000 изображений, чтобы учесть разнообразие условий, освещения и углов съёмки. Это особенно важно для подобных задач, где требуется отличать объекты со схожими визуальными признаками.
2. **Учет разнообразия в данных:** Если каждый тип объекта представлен в различных условиях (например, освещение, погодные условия), можно использовать чуть меньшее количество изображений, так как увеличенное разнообразие позволяет модели обобщать лучше. В каждом изображении должны быть разнообразные примеры дыма, тумана и облаков при разных условиях. Такой подход помогает модели различать классы объектов даже в новых, незнакомых условиях.
3. **Практические эксперименты для оценки достаточности набора данных:** При ограниченных ресурсах можно начать с базового набора данных, протестировать производительность модели и постепенно увеличивать количество данных в зависимости от результатов. Это позволяет оценивать, насколько добавление данных действительно улучшает точность, и не увеличивать объем выборки без необходимости. Начать рекомендуется с 5 000 изображений, равномерно распределив их между категориями (дым, туман, облака), и оценить точность и полноту модели. Для YOLO рекомендуется ориентироваться на показатели точности и полноты около 85–90% как показатель того, что объем выборки достаточен.
4. **Проверка с помощью статистического и модельного анализа:** Для подтверждения достаточности объема данных можно применять статистические

методы, например, t-тест на показателях производительности на валидационном наборе, чтобы определить, действительно ли добавление данных значительно улучшает точность модели. Это помогает избежать чрезмерного увеличения выборки и сосредоточить усилия на полезном расширении данных, если это необходимо.

Чтобы провести расчёт минимального размера выборки, обратимся к формуле для оценки среднего значения выборки. Мы будем исходить из гипотезы, что у нас имеются 1,5 млн изображений для огня и 1,5 млн изображений для дыма. Однако, чтобы не аннотировать каждое изображение, мы постараемся определить минимальное количество аннотированных изображений, достаточное для достижения необходимого качества работы модели.

Формула для расчета размера выборки:

Формула для расчета необходимого объема выборки выглядит следующим образом:

$$n = \frac{Z^2 \cdot p \cdot (1 - p)}{H^2}$$

Где:

n- размер выборки,

Z - Z-оценка для заданного уровня доверия (например, для 95% доверия $Z = 1.96$),

p - предполагаемая доля признака в выборке. Обычно используется максимальное значение 0.5, так как это обеспечивает наибольшую дисперсию и, следовательно, наибольший размер выборки.

Пример расчета:

Параметры:

1. **Уровень доверия:** 95% ($Z = 1.96$)
2. **Допустимая ошибка (H):** Например, возьмем 1% (0.01), чтобы данные были точными.
3. **Доля признака (p):** $p = 0.5$

Подставляем значения в формулу:

$$n = \frac{(1.96)^2 \cdot 0.5 \cdot (1 - 0.5)}{(0.01)^2}$$

Результат:

$$n \approx 9604$$

Таким образом, по результатам расчетов, потребуется **9604 аннотированных изображения** для обеспечения необходимого качества и надежности модели YOLO v8 при условии 95% доверительного интервала и допустимой ошибки в 1%.

Расчет проведен для 1 класса, таким образом, учитывая классы облака, дым, туман, получается $9604 \cdot 3 = 28812$ изображений для обучающей выборки. Это обеспечит нашей модели достаточное количество данных для обучения и улучшит ее точность в распознавании объектов.

Работа с видеомодулем на дроне **ТОРОТЕК DHU290G609**

Описание видеомодуля:

В данном проекте для сбора данных использовался видеомодуль типа **ТОРОТЕК DHU290G609**, устанавливаемый на дрон. Модуль оснащен двумя видеопотоками: **VIS** (видимый спектр) и **IR** (инфракрасный спектр), которые сжимаются с использованием кодека **H264**. Это сжатие позволяет эффективно передавать данные в реальном времени, снижая нагрузку на каналы связи и обеспечивая быструю передачу изображений.

Декодирование видеопотоков:

Для эффективной работы с данными, переданными с дрона, необходимо выполнить процесс декодирования видеопотоков. На стороне наземной станции данные в формате H264 будут подвергаться декодированию до компонентного видеосигнала **RGB** для дальнейшего анализа и обработки. Для этого используется стандартное программное обеспечение для декодирования видеопотоков, которое восстанавливает изображения, обеспечивая точность и качество при преобразовании сжатого видео в формат, пригодный для использования в алгоритмах машинного обучения.

Процесс декодирования включает в себя следующие этапы:

1. **Декодирование H264 потока** — преобразование сжатых видеопотоков (VIS и IR) в необработанные кадровые данные.
2. **Преобразование в RGB** — преобразование декодированных данных в компонентный видеосигнал RGB для дальнейшей работы с изображениями.
3. **Интеграция данных** — комбинация полученных видеопотоков (VIS и IR) для создания многоканальной картины для улучшения точности распознавания объектов, таких как дым, туман и облака.

Передача данных:

Передача видеопотоков между дроном и наземной станцией будет осуществляться по беспроводной связи с использованием сети **5G**, что обеспечит низкую задержку и высокую скорость передачи данных. Это позволяет проводить

анализ в реальном времени и эффективно обрабатывать изображения с высокой частотой.

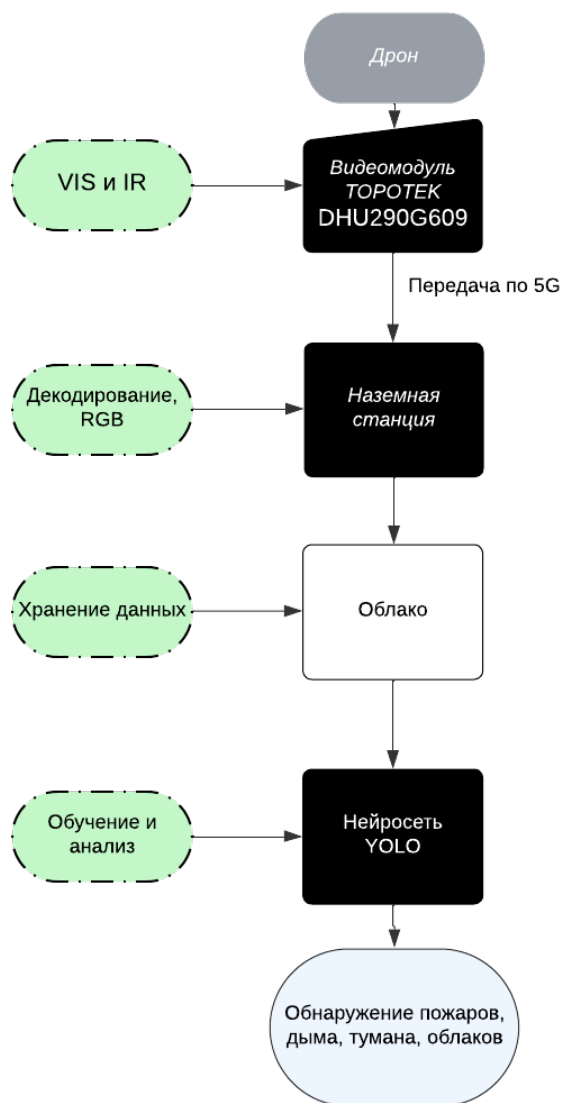


Рис. 2 - Блок-схема работы с данными

Преимущества использования этого подхода:

- **Многоканальная информация:** Комбинированное использование видеопотоков в видимом и инфракрасном спектре позволяет повысить точность распознавания объектов, особенно в сложных погодных условиях, таких как дым, туман или облака.
- **Реальное время:** Использование 5G для передачи данных позволяет осуществлять обработку и анализ видеопотоков в реальном времени, что

критично для оперативного обнаружения и мониторинга природных катастроф, таких как лесные пожары.

Ссылки

1. Расчет объема выборки - <https://neiros.ru/tools/sample-size/>
2. Расчет объема выборки - <https://mobidev.biz/blog/object-detection-small-datasets-use-cases-machine-learning>
3. Расчет объема выборки - <https://blog.roboflow.com/how-to-train-yolov8-on-a-custom-dataset/>
4. Датасет - https://github.com/XiwenChen-Clemson/Flame_2_dataset
5. "Кинетические параметры аэрозолей дыма при горении модифицированной древесины», Е. Н. Покровская, Ф. А. Портнов, 2017
6. Talukdar, Jonti & Gupta, S. & Rajpura, P. & Hegde, Ravi. (2018). Transfer Learning for Object Detection using State-of-the-Art Deep Neural Networks. 78-83. 10.1109/SPIN.2018.8474198.
7. Chawda, Ajay & Vierling, Axel & Berns, Karsten. (2022). Dimensionality of datasets in object detection networks. 10.48550/arXiv.2210.07049.
8. «Anomaly detection in static networks using egonets.» Srijan Sengupta - 2018 - <https://arxiv.org/pdf/1807.08925>
9. «A MEAN FIELD GAME PRICE MODEL WITH NOISE». Diogo Gomes, Julian Gutierrez, Ricardo Ribeiro. - 2020 - <https://arxiv.org/pdf/2003.01945>