Двадцать четвертая международная ежегодная конференция
«Современные проблемы дистанционного зондирования Земли из космоса»

Использование сверточной нейросети для детекции дыма на двухканальных спутниковых снимках

Вербицкий А. В., Мартинов А. О., Ломако А. А., Научно-исследовательское учреждение “Институт прикладных физических проблем имени А. Н. Севченко” Белорусского государственного университета

Актуальность, цель и задачи

Актуальность

- Лесные пожары наносят значительный экологический и экономический ущерб, поэтому важно выявлять их как можно раньше.
- Дымовой шлейф можно заметить на спутниковом снимке ещё до того, как очаг пожара удаётся уверенно определить по другим признакам.
- В малых спутниках формата CubeSat число спектральных каналов может быть ограничено двумя. Поэтому рассматривается метод, который использует только два канала и дополняет основной тепловизионный модуль.

Цель

Оценить возможности двухканального анализа спутниковых изображений для обнаружения дымовых шлейфов и обосновать выбор спектральных каналов. для разрабатываемой съёмочной системы наноспутника.

Задачи

- Обучить классификатор ResNet50 на наборе USTC_SmokeRS и проверить, различает ли он дым и похожие явления.
- Обучить детектор YOLOv8 на спутниковой части набора FASDD_RS, чтобы определять расположение дыма.
- Сравнить три пары каналов по единым для каждой задачи метрикам и изучить типичные ошибки.

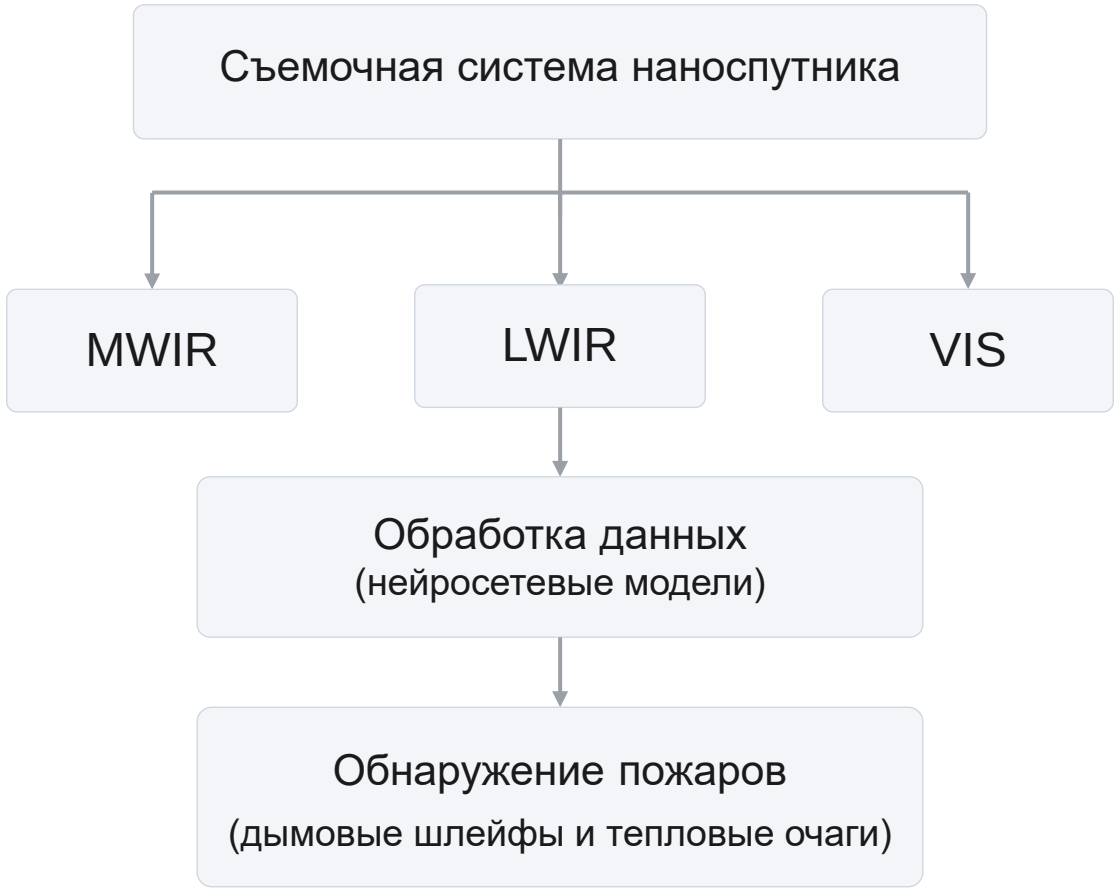
Концепция съёмочной системы наноспутника

Разрабатываемая съёмочная система предназначена для наноспутника размером до 12U и ориентирована на обнаружение лесных пожаров.

Предполагается использование трех сенсоров:

- **MWIR** — обнаружение теплового источника/очага;
- **LWIR** — наблюдение поверхности сквозь дым, оценка теплового контраста;
- **VIS (2 канала)** — обнаружение дымового шлейфа как косвенное подтверждение.

Видимый модуль предполагается реализовать с использованием двух спектральных каналов. Для разделения диапазонов рассматривается применение сдвоенного фильтра с напылением.



Спектральные особенности дыма

Видимый диапазон (VIS)

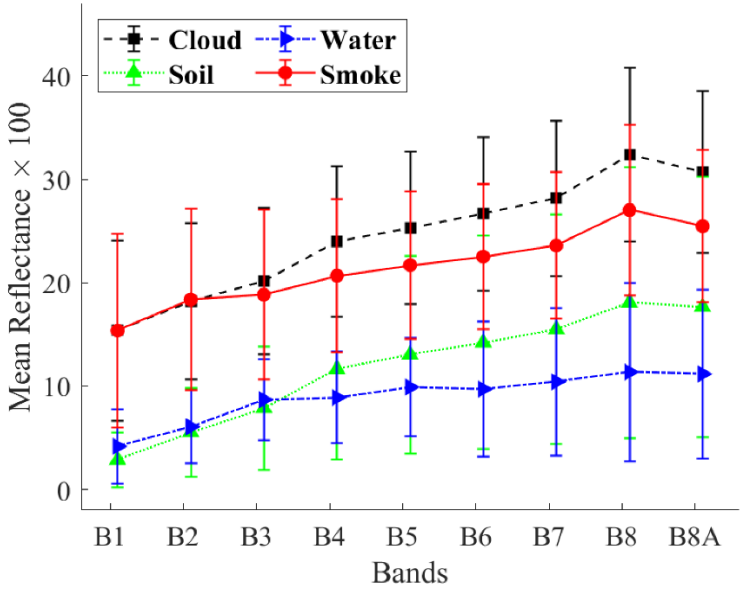
- Дым рассеивает коротковолновое излучение, поэтому в синем канале может выглядеть более ярким, а в красном диапазоне более темным.
- Сопоставление синего и красного каналов позволяет получить дополнительный контраст дымового шлейфа.
- Основная сложность — сходство спектральных характеристик дыма и облаков, особенно в синем и зелёном диапазонах.

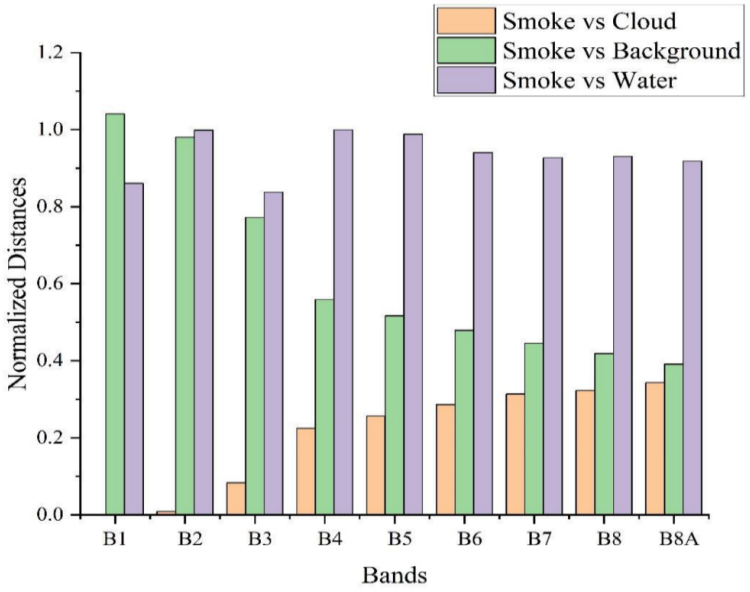
Инфракрасный диапазон (IR)

- Ближний ИК (NIR) позволяет получить дополнительную информацию об отражательных свойствах дыма, облаков и подстилающей поверхности.
- Отдельные полосы NIR могут повышать различимость дыма и облаков.
- Средневолновый и длинноволновый ИК (MWIR и LWIR) используются прежде всего для регистрации теплового излучения очагов пожара.

Спектральные каналы: что известно из литературы

Канал	Что регистрирует	Информативность для дыма	Доступность на малом КА
Blue (450–500 нм)	Коротковолновое отражённое излучение	Полезен для выявления дымового шлейфа за счёт рассеяния	Относительно простая реализация
Red (650–670 нм)	Красную область отражённого излучения	Дополняет Blue, создавая межканальный контраст	Относительно простая реализация
Red Edge (740–790 нм)	Переходную область между красным и NIR	Может давать дополнительную информацию о фоне и растительности	Возможна реализация узкополосного фильтра
NIR (830–880 нм)	Ближнее инфракрасное отражённое излучение	Перспективен для разделения дыма и облаков	Требует отдельного подбора фильтра и оценки чувствительности детектора





Ближний инфракрасный диапазон (NIR), в частности область около 865 нм, представляет интерес для разделения дыма и облаков. Однако окончательный выбор канала должен учитывать различимость всех рассматриваемых объектов и характеристики съёмочной аппаратуры.

В данной работе NIR экспериментально не проверялся из-за отсутствия подходящих открытых наборов данных. Поэтому для сравнения доступных двухканальных комбинаций использовался красный канал.

2023 - Detection of Smoke from Straw Burning Using Sentinel-2 Satellite Data and an Improved YOLOv5s Algorithm https://www.mdpi.com/2072-4292/15/10/2641

Каналы и их комбинации

Спектральные каналы

Канал	MODIS	Sentinel-2, Landsat-8
Blue	459–479 нм	490 нм
Green	545–565 нм	560 нм
Red	620–670 нм	665 нм

Зелёный канал помогает учитывать особенности растительного покрова, но сам по себе не всегда позволяет уверенно отличить дым от облаков.

Двухканальные комбинации

$K_1 = (Blue, Red)$

$K_2 = (Green, Red)$

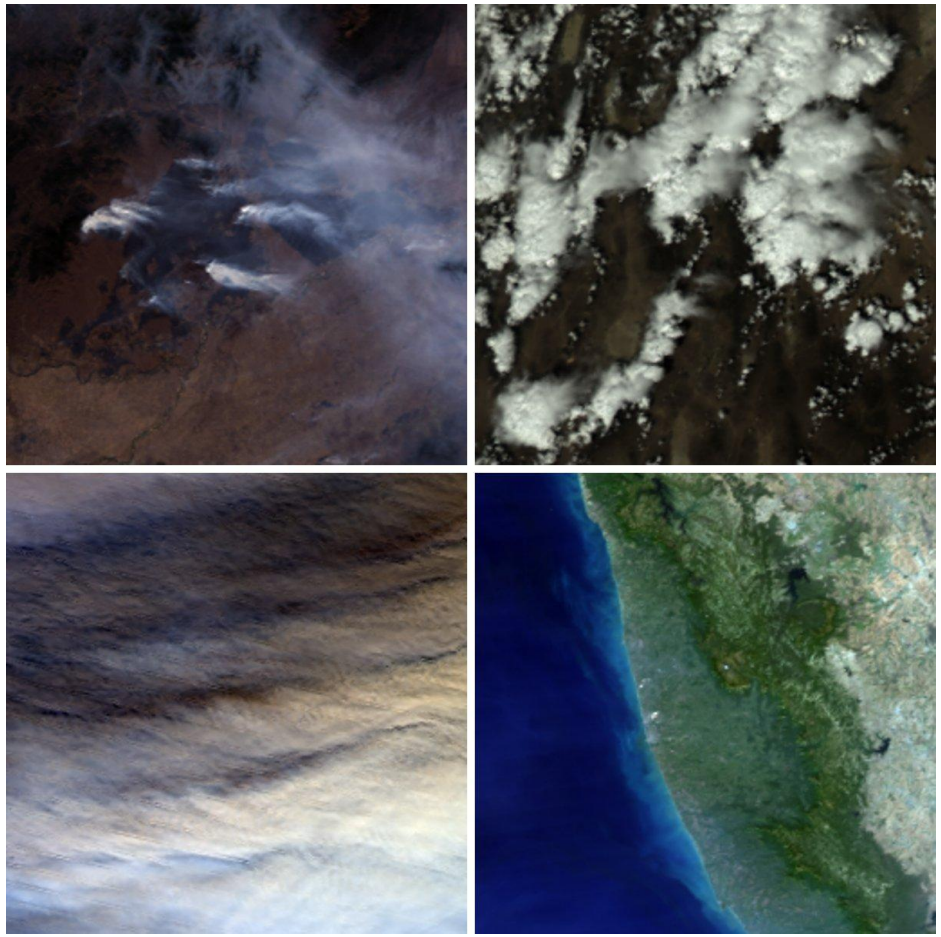
$K_3 = ((Green + Blue)/2, Red)$

Данные: два набора

Характеристика	USTC_SmokeRS	FASDD_RS
Источник	MODIS (спектрорадиометр)	Sentinel-2, Landsat-8
Количество снимков	6225	2223
Размер снимка	256×256 пикселей	1000×1000 пикселей
Разрешение	1 км/пиксель	10 и 30 м/пиксель
Формат	.parquet	.tif
Задача	Классификация (6 классов)	Детекция дыма
Модель	ResNet50	YOLOv8

Использованы два набора, поскольку они содержат спутниковые изображения, подходящие для разных этапов задачи. USTC_SmokeRS применяется для классификации всего снимка, а FASDD_RS — для определения положения дымовых областей.

Примеры снимков



USTC_SmokeRS (MODIS): cloud, dust, haze, land, seaside, smoke



FASDD_RS (Sentinel-2, Landsat-8)

Разметка обучающих данных

USTC_SmokeRS — классификация

Каждому снимку присвоена одна метка класса. Модель обучается определять, к какой категории относится изображение целиком.

6 классов:

- cloud — облака
- dust — пыль
- haze — дымка
- land — суша
- seaside — побережье
- smoke — дым

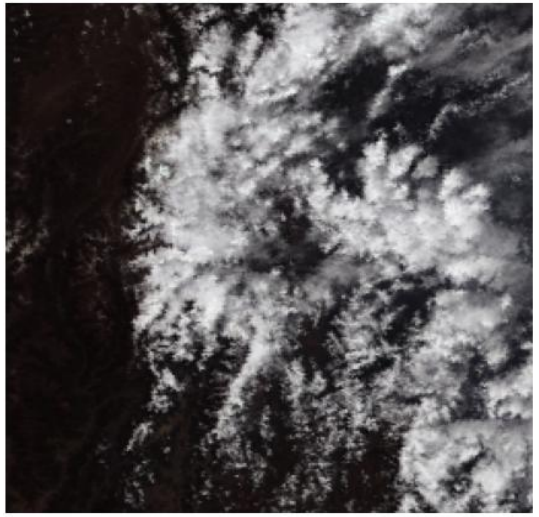
FASDD_RS — детекция дыма

Для спутниковых изображений Sentinel-2 и Landsat-8 используется эталонная разметка, указывающая области, которые относятся к дыму.

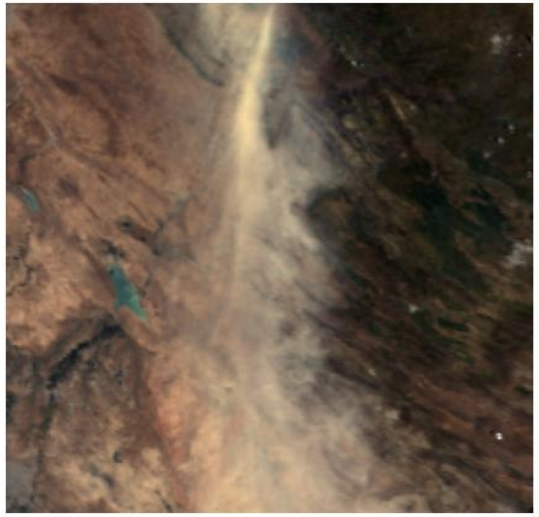
Во время обучения предсказания модели сравниваются с эталоном. Это помогает оценить, находит ли сеть дым в правильном месте.

Примеры разметки данных (USTC_SmokeRS)

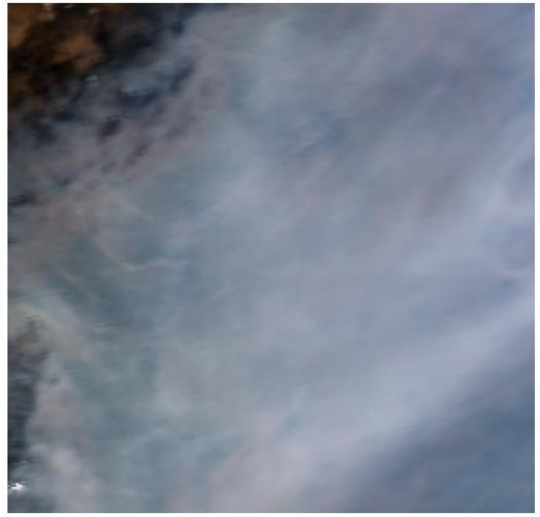
Класс 0: Cloud



Класс 1: Dust



Класс 2: Haze



Класс 3: Land



Класс 4: Seaside

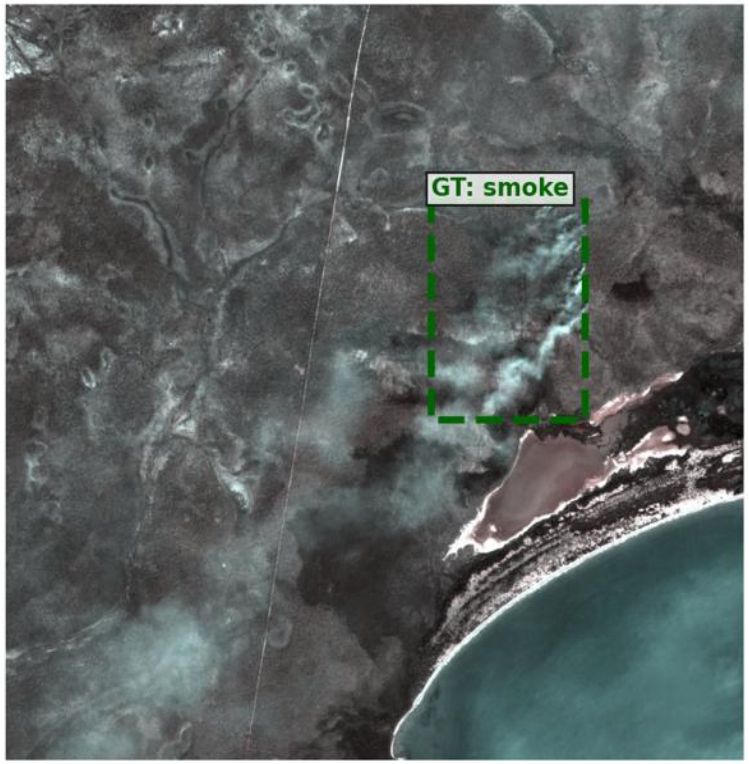


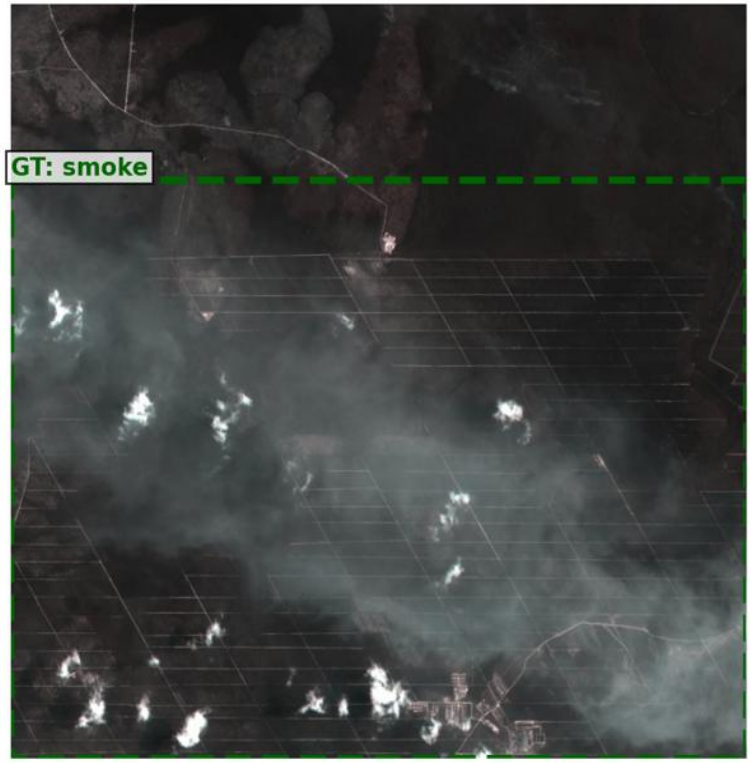
Класс 5: Smoke

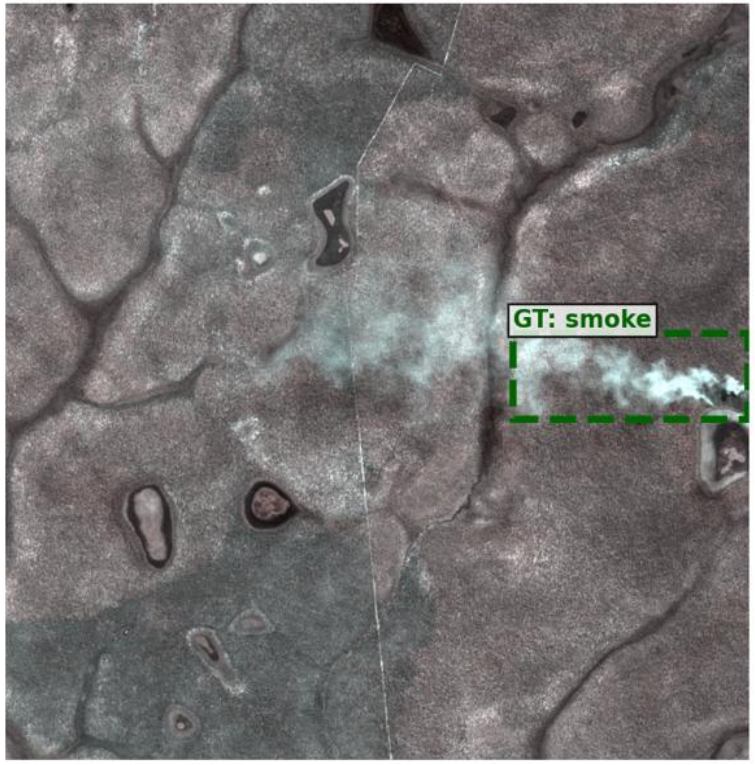


Разметка в датасете USTC_SmokeRS: изображение — класс (число 0-5).

Примеры разметки данных (FASDD_RS)







Разметка в датасете FASDD_RS: изображению с дымом соответствует маска в виде прямоугольника.

Классификация: постановка задачи

Вход

На вход поступает снимок MODIS размером 256×256 пикселей с пространственным разрешением около 1 км на пиксель. Из него выбираются два спектральных канала; затем изображение приводится к размеру 224×224 пикселя для подачи в нейросеть.

Выход

На выходе модель присваивает всему снимку один из шести классов: облака (cloud), пыль (dust), дымка (haze), суша (land), побережье (seaside) или дым (smoke). На этом этапе определяется наличие дыма на снимке, но не его точные границы.

Класс	cloud	dust	haze	land	seaside	smoke
Снимков	1164	1009	1002	1027	1007	1016

Разбиение: обучение 64% (3984), валидация 16% (999), тест 20% (1242), случайное.

Предобработка снимков

Этап	Операции	Назначение
Обучение	Уменьшение 256→230, случайная обрезка до 224×224, случайные отражения по горизонтали и вертикали, нормализация (среднее 0,5; СКО 0,5)	Аугментация. Малые области дыма на краях снимка не теряются
Валидация, тест	Прямое уменьшение до 224×224, та же нормализация	Детерминированная оценка

Нормализация приводит значения пикселей к единому масштабу, чтобы сети было проще обучаться. В этой работе использовано среднее 0,5 и стандартное отклонение 0,5.

Архитектура ResNet50

Блок	Выход	Каналы
Вход	224×224	2
Conv 7×7, pool	56×56	64
Стадия 1 (3 блока)	56×56	256
Стадия 2 (4 блока)	28×28	512
Стадия 3 (6 блоков)	14×14	1024
Стадия 4 (3 блока)	7×7	2048
Global avg pool	1×1	2048
FC → 6 классов	6	—

- ResNet50 — свёрточная нейросеть из 50 слоёв с обучаемыми параметрами. В начале она выделяет простые признаки изображения, например границы и текстуры; более глубокие слои объединяют их в сложные формы и структуры, характерные для дыма.
- Затем усреднение по пространственным областям превращает карту признаков в вектор из 2048 чисел. Финальный слой использует этот вектор, чтобы оценить вероятность каждого из шести классов.

Остаточный блок

$$y = F(x) + x$$

1×1 conv
Уменьшает число каналов

3×3 conv
Извлекает пространственные признаки

1×1 conv
Восстанавливает число каналов

Сумма + ReLU
Складывает результат F(x) со входом x

- В остаточном блоке исходные признаки x передаются по короткому соединению и складываются с результатом преобразования F(x). Поэтому блоку не обязательно заново строить всё представление: он может выучить только необходимое изменение.
- Такое соединение помогает обучать глубокую сеть, поскольку сигнал обучения может проходить по более короткому пути. Последовательность свёрток 1×1 → 3×3 → 1×1 также снижает вычислительные затраты.

Обучение классификатора

Взвешенная функция потерь

$$L = -\Sigma\, w_c \cdot y_c \cdot \log p_c$$

- Класс smoke получил вес 1,5, а остальные классы — вес 1. Это означает, что ошибка при распознавании дыма сильнее влияет на функцию потерь. Такой приём нужен, чтобы модель уделяла больше внимания главному для мониторинга классу, хотя он может увеличить число ложных тревог.

Параметр	Значение
Batch size	32
Эпох	30
Размер входа	224
Learning rate	0,001
Шаг планировщика / gamma	30 / 0,1
Видеокарта	RTX 3050, 8 ГБ

Метрики классификации

$$Precision = TP / (TP + FP)$$

$$Recall = TP / (TP + FN)$$

$$F1 = 2 \cdot P \cdot R / (P + R)$$

- TP — случай, когда дым распознан правильно; FP — ложная тревога, например облако принято за дым; FN — реальный дым, который модель пропустила.
- Precision показывает, какая доля всех срабатываний «дым» оказалась правильной. Recall показывает, какую долю реального дыма удалось найти. F1 объединяет эти две характеристики и помогает оценить их баланс. Все три метрики здесь рассчитаны для класса smoke.

Результаты классификации (класс smoke)

Модель	Precision	Recall	F1	Общая точность
Blue + Red, вес 1	71,79%	73,20%	72,49%	81,69%
Blue + Red, вес 1,5	78,71%	79,74%	79,22%	83,19%
Green + Red	63,45%	81,70%	71,43%	79,66%
(Green+Blue)/2 + Red	83,46%	72,55%	77,62%	81,15%

- Лучший баланс Precision и Recall: Blue + Red с весом smoke 1,5.
- Вес класса поднял F1 с 72,49% до 79,22%.
- Green + Red: высокий Recall, но много ложных тревог.

Влияние веса класса: точность по классам

Класс	Вес 1	Вес 1,5	Изменение
cloud	86,86%	78,86%	−8,00
dust	90,07%	81,46%	−8,61
haze	66,00%	79,33%	+13,33
land	87,66%	85,71%	−1,95
seaside	85,43%	94,70%	+9,27
smoke	73,20%	79,74%	+6,54

Вес улучшился smoke, haze и seaside, но ухудшилось cloud и dust. Общая точность выросла с 81,69% до 83,19%.

Детекция: постановка задачи

Вход

На вход подаётся снимок Sentinel-2 или Landsat-8 размером 1000×1000 пикселей. Разрешение — 10 или 30 м на пиксель. Используются два спектральных канала.

Выход

Модель выделяет рамками области, похожие на дым, и оценивает уверенность каждого обнаружения. Так определяется не только наличие дыма, но и его примерное положение.

Эталон

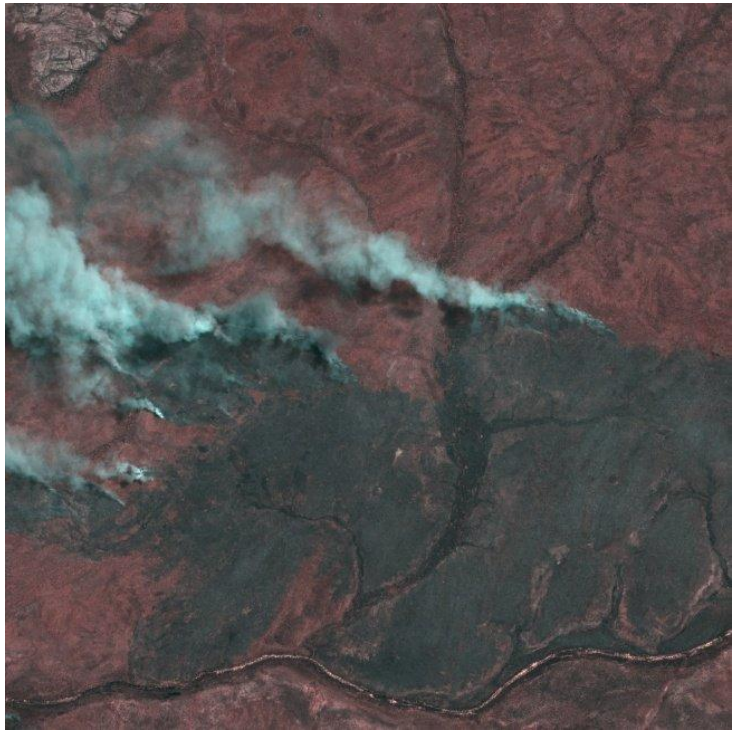
Для обучения используются маски (bounding boxes) из FASDD_RS. Маска — это разметка, показывающая, где на снимке находится дым.

- Классификация отвечает на вопрос «есть ли дым на снимке», детекция — «где именно».
- YOLOv8 выбрана для детекции объектов: есть готовые данные и маски.

Подготовка данных для YOLOv8



Blue + Red + Red



(Green+Blue)/2 + Red + Red

- Сначала выбираются два спектральных канала. Затем они преобразуются в JPG и формируются как изображение из трёх каналов: второй канал записывается дважды, поскольку стандартный вход YOLOv8 рассчитан на три канала.
- Значения пикселей приводятся к диапазону 0–255, изображение масштабируется до 640×640 пикселей. Дублирование не создаёт новых спектральных данных — оно лишь позволяет использовать стандартную архитектуру модели.

Архитектура YOLOv8

Backbone

Извлекает признаки изображения с помощью свёрток и специальных блоков. На этом этапе сеть обнаруживает границы, текстуры и возможные структуры дымового шлейфа.

Neck

Объединяет признаки разного масштаба. Это важно, потому что на спутниковом снимке дымовой шлейф может занимать как небольшую, так и большую область.

Head

Формирует предсказания: где находится объект и насколько модель уверена, что это дым. Сеть не использует заранее заданные опорные рамки.

Масштаб	Сетка (вход 640)	Шаг	Что находит
P3	80×80	8	мелкие области дыма
P4	40×40	16	средние
P5	20×20	32	крупные шлейфы

YOLOv8 — одностадийный детектор: он предсказывает положения объектов за один проход по изображению. Если несколько рамок описывают одну и ту же область, алгоритм NMS оставляет наиболее подходящие варианты.

Метрики детекции

$$IoU = |A \cap B| / |A \cup B|$$

- IoU показывает, насколько предсказанная область совпадает с эталонной разметкой: чем больше перекрытие, тем лучше совпадение.
- mAP50 оценивает обнаружение при условии, что перекрытие составляет не менее 50%. mAP50–95 усредняет оценку для более строгих порогов от 50 до 95% и поэтому дополнительно отражает точность положения рамок.

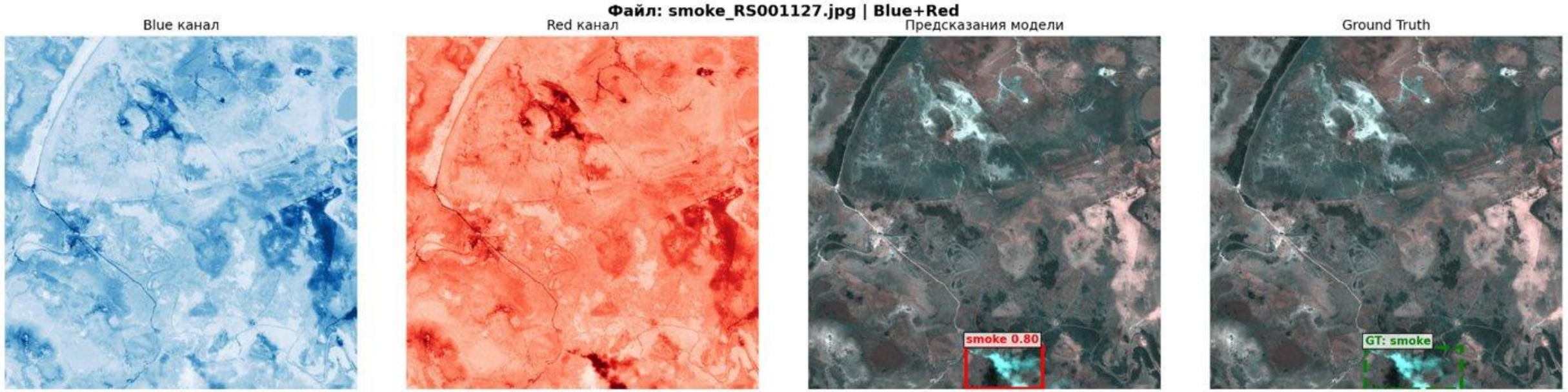
Параметр	Значение
Batch size	32
Эпох	до 50
Размер входа	640
Learning rate	0,01
Patience	15
Workers	4

Результаты детекции

Модель	mAP50	mAP50-95	Precision	Recall
Blue + Red	0,6973	0,4034	0,7272	0,6469
Green + Red	0,5820	0,3338	0,6970	0,6322
(Green+Blue)/2 + Red	0,6060	0,3609	0,6629	0,6448

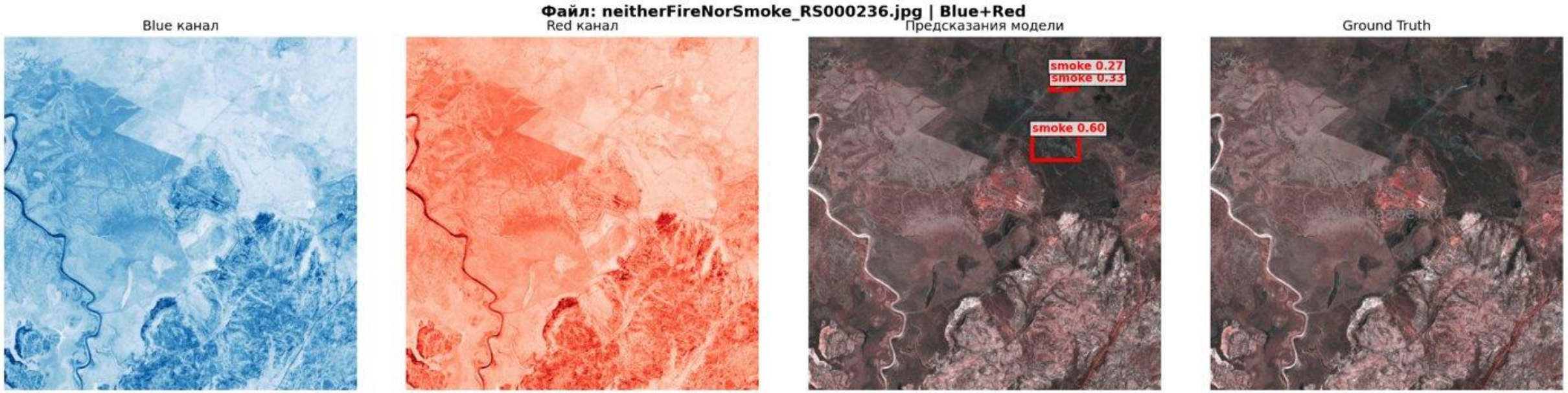
- Комбинация Blue + Red показала лучшие значения всех рассмотренных метрик среди трёх вариантов.
- При интерпретации результатов важно учитывать качество эталонной разметки: маска может не охватывать весь дымовой шлейф или включать участки, которые к дыму не относятся. В таких случаях даже визуально разумное предсказание может получить заниженную оценку.

Пример верной детекции (Blue + Red)



Слева направо: канал Blue, канал Red, предсказание модели, эталонная маска.
Модель находит дым, в ряде случаев точнее эталонной маски.

Ложное срабатывание: облака



На снимке нет заметного дыма или огня, однако модель выделяет облака и другие похожие области как возможный дым. Это пример ложного срабатывания.

В синем и зелёном каналах облака могут выглядеть так же, как дым. В красном канале дым чаще выглядит темнее облаков, но это различие не универсально. Поэтому двух каналов не всегда достаточно, чтобы устранить все ложные срабатывания.

Сравнение комбинаций каналов

Комбинация	Классификация, F1	Детекция, mAP50
Blue + Red	79,22%	0,6973
Green + Red	71,43%	0,5820
(Green+Blue)/2 + Red	77,62%	0,6060

Blue + Red лучшая комбинация среди проверенных пар каналов, Green + Red в обеих худшая. Результат согласуется с физикой дыма: контраст Blue и Red максимален.

Ограничения и перспективы

Ограничения

- Облака и дым могут иметь сходный вид, что приводит к ложным срабатываниям.
- Оценка детекции зависит от полноты и точности эталонных масок.
- Классификация и детекция проверялись на разных наборах данных и спутниковых системах, поэтому результаты этих двух этапов нельзя напрямую считать полностью сопоставимыми.

Планы

- Расширить обучающие данные снимками облаков и других объектов, похожих на дым.
- Подобрать параметры обучения и веса классов.
- При появлении подходящих открытых данных проверить ближний инфракрасный канал (NIR).
- Рассматривать детектор как дополнительный инструмент к тепловизионному модулю, а не как его замену.

Заключение

В работе была показана возможность использования двухканальных спутниковых изображений для обнаружения дыма с помощью нейросетевых моделей. Среди проверенных комбинаций видимого диапазона лучшие результаты показала пара Blue и Red. Полученные результаты рассматриваются как основа для дальнейшей разработки оптического модуля съёмочной системы наноспутника. При этом отдельный интерес представляет ближний инфракрасный диапазон, который, согласно литературным данным, может улучшить разделение дыма и облаков. Его экспериментальная проверка станет одним из следующих этапов работы.

Спасибо за внимание!



Институт прикладных физических проблем им. А.Н.Севченко Белорусского государственного университета,
Минск, Республика Беларусь.

Отдел аэрокосмических исследований

Телефон: +375 (17) 396 44 09
Электронная почта: remsens@mail.ru
Сайт: http://remsens.by