

Команда УИИ

Кейс

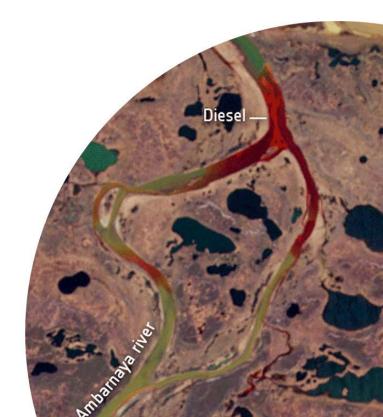
Разработка системы для обнаружения, оценки и мониторинга нефтеразливов (экологического ущерба) с использованием технологий дистанционного зондирования Земли.

Проблема

Региональным органам исполнительной власти из-за нехватки специалистов сложно оперативно обнаруживать аварии, оценивать нанесенный природе ущерб, а также собирать материалы по каждой из них для привлечения виновных лиц к ответственности.

Решение

Создан веб-сервис для анализа спутниковых снимков с целью выявления загрязнений окружающей среды. На основе программного кода производится очистка входных данных и последующее соответствие снимков с координатами, видом загрязнения, датой снимка. Особенность данного сервиса является простое и интуитивное управление не требующее знаний программирования.





РАБОТА С ДАННЫМИ

- Чистка данных (ненужные столбцы, отсутствующие файлы, отсутствующие bbox)
- Создание датафреймов в необходимом формате для работы модели
- Совмещение данных со снимков с таблицей загрязнений в одном датафрейме.
- Аугментация данных. В методы аугментации входят: перемешивание экземпляров фото по каждому квадрату, зашумление картинки, случайные скачки контрастности, вращение, случайная обрезка, отражение, транспонирование и смещение с отзеркаливанием.





НЕЙРОННАЯ СЕТЬ

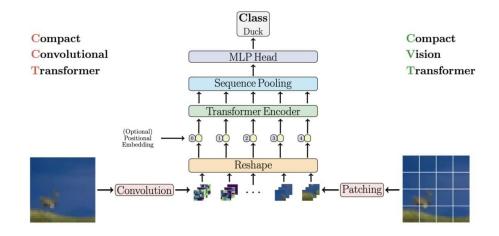
Новые подходы

Использование модели трансформера (ССТ)

Последнее время трансформеры показывают впечатляющие результаты не только в области NLP, но и в области CV. Модель показала высокие результаты по скорости обработки и точности на известном датасете cifar, кроме того модель имеет мало параметров, малый вес, и высокую скорость работы.

Использование Stochastic depth

Stochastic depth - техника регуляризации, которая случайным образом блокирует работу некоторых слоев нейросети. По смыслу близка к "дропауту" - Dropout, с той разницей, что дропаут блокирует отдельные нейроны.

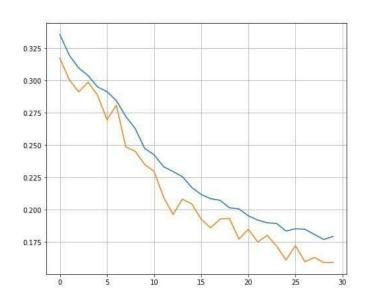


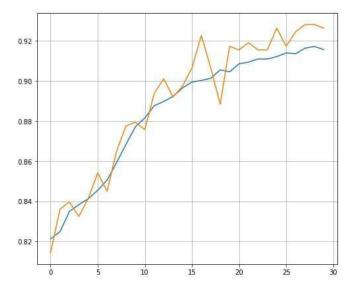
Трансформер делит картинку на патчи - меньшие части, анализирует их и дает им токены значения содержимого, и токены значения позиции, а дальше работает как в нлп - анализирует, как коррелируют значения содержимого каждого кусочка картинки с тем, на каком месте эта часть картинки находится.



ТОЧНОСТЬ

Ошибки и точности нейросети во время первых 30 эпох обучения

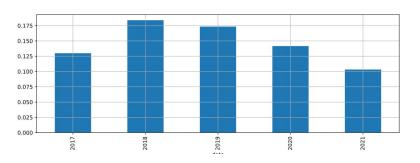


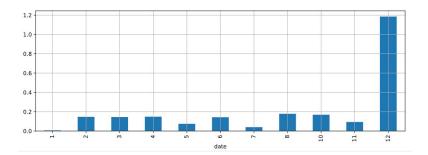




АНАЛИЗ ДАННЫХ

На графиках видно, что наибольшие загрязнения произошли в 2018 году и наибольшее число загрязнений приходится на декабрь.









ИНТЕРПРЕТАЦИЯ МОДЕЛИ И АНАЛИТИКА

Используемые материалы:

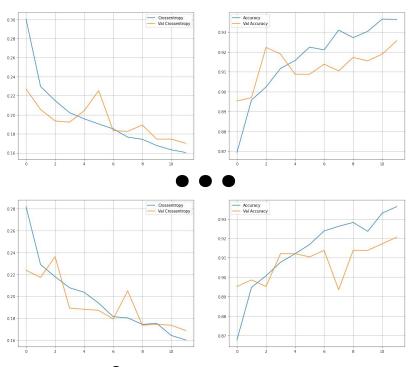
- OIL SPILL MONITORING HANDBOOK
- SENTINEL L2A Product Definition Document

Интерпретируемость модели

Поочередное удаление каналов Определение результата точности моделей после удаления каналов.

В результате можно заметить, что удаление большинства каналов не особо сказалось на точности модели. Однако удаление каналов 3го, 8го привело к более заметному падению точности.

Удаляем каналы по очереди



Смотрим на точность Максимальное влияние 3 и 8 канал



ПОЛЬЗОВАТЕЛЬСКИЙ ИНТЕРФЕЙС



Новаторские идеи и фичи

- 1. Распаковка датасета
- 2. Обработка загруженного изображения нейросетью с выдачей и генерацией интерактивной карты
- 3. Просмотр снимков на карте
- 4. Вероятность разлива
- 5. Экспорт в PDF





стоимость и сроки

Внедрение

Для запуска пилотной версии с базовым функционалом – **3 месяца**

Стоимость внедрения 960 000 руб.

Команда готова к внедрению решения



КОМАНДА УИИ



Качалкин Артём

Data Scientist



Соснин Дмитрий

Data Scientist



Курочкин Алесей

Data Scientist



Хуторной Борис

Data Scientist



Домненко Алексей

Data Scientist Full-stack



ПОЛЕЗНЫЕ ССЫЛКИ

Пользовательский интерфейс http://144.76.152.13:9999/system/

логин: admin

пароль: предоставлен экспертам

Видеодемонстрация



