

Автоматическая типизация горных пород

Александр Смирнов

науч. рук.: ст. преп. Смирнов Михаил Николаевич

27.12.2019г

Введение

- ▶ Способ разведки нефтяных месторождений — разведка буром
 - ▶ Во время бурения извлекают керн
 - ▶ Видно, как залегают пласты породы
 - ▶ Позволяют обнаружить породы-коллекторы, оценить их емкостные и фильтрационные свойства
- ▶ Керн — цилиндрический монолит горной породы, получаемый путём кольцевого разрушения забоя скважин при бурении

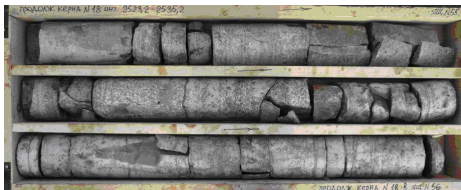


Рис.: Пример керна

Введение (2)

- ▶ Фотографии керн в ультрафиолетовом свете
 - ▶ Позволяют выделить в разрезе нефтенасыщенные участки
 - ▶ Позволяют выделить текстурные характеристики, связанные с особенностями условий осадконакопления пород

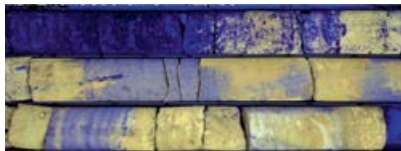


Рис.: Фотографии керн в ультрафиолетовом свете. Неравномерное желтое свечение – неравномерно нефтенасыщенный песчаник.

Проблема

- ▶ Информация о керне описывается послойно: один слой - один тип породы
- ▶ Можно размечать изображения керна с большей точностью
 - ▶ Более точные модели пластов

Цели

- ▶ Получение описания керна на основе выборки фотографий
 - ▶ Тип породы с точностью до 20 см.
 - ▶ Карбонатность с точностью до 10 см.
 - ▶ Нефтенасыщенность с точностью до 10 см.
 - ▶ Разрушенность с точностью до 5 см.
- ▶ Написание удобной для пользователя обёртки над полученным решением для последующего использования

Задачи

- ▶ Произвести разведочный анализ предоставленных данных
- ▶ Ознакомиться с возможными решениями
- ▶ Реализовать решения и найти лучшие
- ▶ Сравнить результаты с уже имеющимися у заказчика
- ▶ Создать оболочку для удобного использования решения

Исходные данные

- ▶ Фотографии керна
- ▶ Те же самые фотографии керна, но в ультрафиолетовом освещении

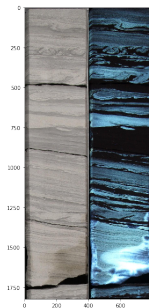


Рис.: Пример из исходных данных. Слева — фото керна, справа — фото того же керна, но в УФ.

Исходные данные (2)

- ▶ Таблица с описанием каждой фотографии
- ▶ Необходимо предсказывать:
 - ▶ Rock (тип породы)
 - ▶ Carbonate (карбонатность)
 - ▶ Ruin (разрушенность)
 - ▶ Saturation (нефтенасыщенность)

	0	1
Id	1000000	1000001
PhotoTop	0	0
PhotoDown	1	1
Rock	песчаник	песчаник
Carbonate	не карбонатный	не карбонатный
Ruin	не разрушен	не разрушен
Saturation	нефтенасыщенные	нефтенасыщенные

Таблица: Пример из таблицы исходных данных. Предоставлена информация о первых двух записях — один и тот же керн, обычная фотография и фотография в УФ.

Анализ данных

- ▶ Поля **PhotoUp** и **PhotoDown** означают начало и конец данной фотографии в данном образце керна
- ▶ **PhotoUp**=0 и **PhotoDown**=1 \Rightarrow длина керна на фотографии — 1 метр
- ▶ В категории **Rock** очень много значений, которые имеют малое количество экземпляров в сравнении с другими категориями \Rightarrow выбросим их после согласования с заказчиком

тип породы	количество экземпляров
песчаник	2482
аргиллит	1220
алевролит	1138
переслой	686

Таблица: Распределение категории "тип породы".

Анализ подходов

- ▶ Классификация изображений
- ▶ Распознавание образов на изображениях
- ▶ Сегментация изображений

Анализ подходов (2)

- ▶ Задача классификации — определяем входное изображение в один из классов
- ▶ Методы машинного обучения
 - ▶ Выделяем признаки с помощью алгоритмов
 - ▶ Гистограмма направленных градиентов
 - ▶ Метод Виолы — Джонса
 - ▶ Используем техники классификации
 - ▶ SVM
- ▶ Методы глубокого обучения
 - ▶ Свёрточные нейронные сети
- ▶ Лучшие решения — свёрточные нейронные сети

Анализ подходов (3)

- ▶ Задача распознавания образов — определяем на фотографии границы объектов известных классов
- ▶ То же, что и классификация, только с помощью алгоритмов сужаем область, в которой находится объект
- ▶ Для более точного распознавания подаём на вход не только фотографию с отмеченной категорией, но и область, в которой находится данный объект

Анализ подходов (4)

- ▶ Сегментация изображений — это процесс присвоения таких меток каждому пикселю изображения, что пиксели с одинаковыми метками имеют общие визуальные характеристики
- ▶ Для достижения большей точности некоторые алгоритмы помимо фотографии принимают изображение с размеченными по категориям пикселями

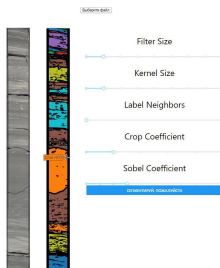


Рис.: Пример работы сегментационного алгоритма.

Решение

- ▶ Будем решать задачу классификации
- ▶ Подготовим данные
- ▶ Выберем алгоритм
- ▶ Сравним результаты

Подготовка данных

- ▶ В каждой категории классификации количество экземпляров не сбалансировано \Rightarrow применяем **oversampling** (увеличение количества экземпляров в недостающих классах путём их копирования)
 - ▶ Алгоритмы не учитывают перевес одних классов над другими в количестве элементов
- ▶ Категории должны быть определены с фиксированной точностью \Rightarrow на вход поступает фотография керна высотой, к примеру, 1 метр, она разделяется на 5 фотографий керна длиной 20 сантиметров каждая
- ▶ Подготовленные подобным образом фотографии будут подаваться на вход ниже описанным алгоритмам

Выбор алгоритмов

- ▶ Свёрточная нейронная сеть.
- ▶ Feature extraction на основе свёрточной нейронной сети + kNN

Свёрточная нейронная сеть

- ▶ Имеются данные двух типов — обычные фотографии и фотографии в ультрафиолетовом свете
- ▶ Будем использовать предобученные модели
- ▶ Возможные решения:
 - ▶ Отдельно обучать алгоритмы на классификацию обычных фотографий и фотографий в УФ. Объединять их выводы.
 - ▶ Имея в сумме 6 каналов цвета из двух трёхканальных фотографий, можно обучать алгоритмы выборочно по 3-м каналам
 - ▶ Обучать $C_6^3 = 20$ алгоритмов на каждую тройку каналов. Объединять их выводы

Свёрточная нейронная сеть (2)

- ▶ Рассмотрим первый подход
- ▶ Протестируем и сравним точность предобученных архитектур
- ▶ Преобразуем наши изображения до формата входных данных обученных сетей

data type	model	epochs	val acc	roc auc
non-UV	VGG16	3	0.75	0.93
non-UV	ResNet50	60	0.69	0.85
non-UV	DenseNet121	221	0.6	0.7
non-UV	Xception	34	0.62	0.74
UV	VGG16	11	0.72	0.88
UV	ResNet50	71	0.62	0.77
UV	DenseNet121	154	0.52	0.64
UV	Xception	60	0.49	0.55

Таблица: Результаты использования предобученных архитектур на примере классификации типа породы.

Свёрточная нейронная сеть (3)

- ▶ Top-1 accuracy вышла 0.75
- ▶ Неплохо, учитывая, что мы используем только половину информации
- ▶ Данный подход может дать ещё большую точность, если эту модель совместить с моделью для классификации по УФ фотографиям

Заключение

- ▶ Что было сделано в рамках данной работы
 - ▶ Изучена предметная область
 - ▶ Проведены изучение и подготовка данных
 - ▶ Проведен разбор возможных решений
 - ▶ Реализовано одно решение
- ▶ Что планируется сделать
 - ▶ Улучшить реализованное решение
 - ▶ Реализовать метрическую классификацию
 - ▶ Реализовать распознавание и сегментацию
 - ▶ Сравнить результаты между собой
 - ▶ Обернуть решение в удобный интерфейс