

Санкт-Петербургский государственный университет

Математическое обеспечение и администрирование информационных
систем

Кафедра информационно-аналитических систем

Смирнов Александр Львович

Автоматическая типизация горных пород

Курсовая работа

Научный руководитель:
ст. преп. Смирнов М. Н.

Санкт-Петербург
2020

Оглавление

Введение	3
1. Цели работы	7
2. Постановка задачи	8
3. Исходные данные	9
3.1. Анализ данных	11
4. Анализ подходов	13
4.1. Классификация	13
4.2. Распознавание образов	13
4.3. Сегментация	14
5. Решение	15
5.1. Подготовка данных	15
5.2. Выбор алгоритмов	15
5.3. Свёрточная нейронная сеть	16
5.4. Метрическая классификация изображения	17
Заключение	18
Список литературы	19

Введение

Существует множество способов разведки нефтяных месторождений. Один из них — разведка буром: во время бурения аккуратно извлекают керн — цилиндрические столбики породы, по которым ясно видно, как залегают пласты. Полученные образцы позволяют обнаружить породы-коллекторы, оценить их емкостные и фильтрационные свойства.

Что такое керн

Керн — цилиндрический монолит горной породы, получаемый путём кольцевого разрушения забоя скважин при бурении.

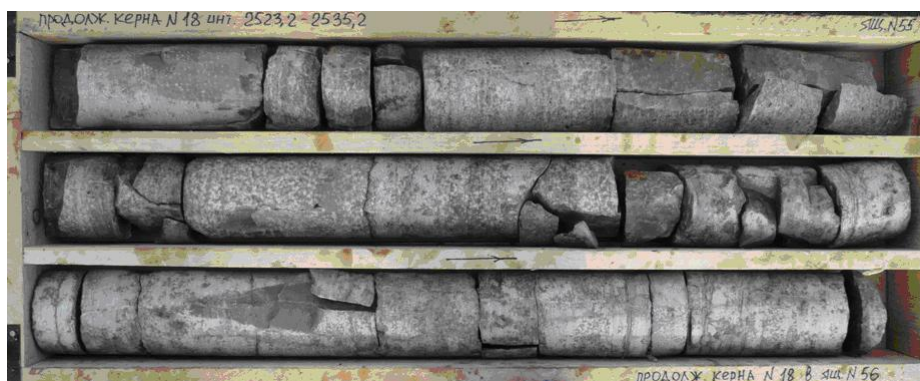


Рис. 1: Пример керна

К сохранности и качеству керна предъявляются требования, обеспечивающие достоверность сведений о составе и строении вскрытых скважиной горной породы и полезных ископаемых. Сохранность керна оценивается его линейным (или объёмным) выходом — процентным отношением суммарной длины (или фактической массы) поднятого керна к длине пробуренного интервала (или расчётной массе для пробуренного интервала) скважины.

В дальнейшем керн исследуется и анализируется (химический, спектральный, петрографический и другие анализы) в лаборатории с помощью различных методов и на различном оборудовании, в зависимости от того, какие данные должны быть получены.

Выход керна по руде обычно колеблется от 50 до 80%. В плотных и однородных рудах и породах он повышается до 100%. В мягких и сильно трещиноватых рудах выход керна иногда снижается до нуля. При отсутствии или малом выходе керна в пробу поступает шлам, вследствие чего качество опробования значительно снижается. Учитывая это, следует добиваться максимального выхода бурового керна.

При опробовании массивных и вкрапленных руд большой мощности применяется секционный отбор проб с длиной керна отдельной пробы 1, 2 или 3 метра, а иногда 5 метров в соответствии с методами предстоящей эксплуатации.

Описание разреза начинается с общего осмотра керна (или его части) и уточнения его местоположения в разрезе скважины. Керна, поднятый и очищенный от бурового раствора, укладывают в специальные кернаые ящики, изготовленные из дерева и разделенные на продольные секции. После этого проводятся исследования состава, выделение маркирующих слоёв и прочее.

Исследование керна

Изучение нефтегазоносных скважин по керну имеет свои специфические особенности. Они заключаются в том, что по керну скважин получают в основном геологическую информацию, связанную с закономерностями вертикального строения разрезов (последовательность и характер напластования, мощность слоев, литологический состав отложений, текстурно-структурные особенности пород и т.д.).

Кроме того, отбор керна в скважинах осуществляется не полностью, поэтому полученные сведения могут носить обрывочный характер и требуют глубокого анализа строения разрезов, вскрытых ранее пробуренными скважинами, и привлечения данных геофизических исследований.

Осадочные толщи имеют слоистое (часто ритмичное) строение и представляют многократное и разномасштабное повторение (чередование) пород. Поэтому при осмотре и описании кернаых колонок, преж-

де всего, выделяются слои – геологические тела, имеющие существенно однородный литологический состав (часто одинаковую окраску), обладающие ясно выраженными подошвой и кровлей и значительной толщиной (мощностью).

Определение наличия нефти

Фотографии керна в ультрафиолетовом свете (Рис. 2) позволяют выделить в разрезе нефтенасыщенные участки, выявить текстурные характеристики, связанные с особенностями условий осадконакопления пород. Нефтенасыщенные интервалы керна светятся в ультрафиолетовом свете в спектре от голубого до буровато-оранжевого цвета. Чем выше плотность углеводородов и насыщенность ими пород, тем больше желтых, оранжевых и коричневых цветов. [3]

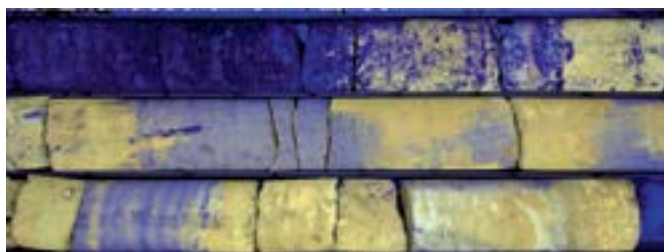


Рис. 2: Фотографии керна в ультрафиолетовом свете. Неравномерное желтое свечение – неравномерно нефтенасыщенный песчаник.

Нефтепроявления могут заключаться в выходах жидкой нефти и подъеме нефтесодержащих пород, в примазках нефти по трещинам в породах, в тонких пленках нефти на воде и т.д. Нефть может вытекать непосредственно из коренных пород, из наносов; может скапливаться в виде толстых плёнок на поверхности воды более или менее далеко от места выхода нефтеносных пород на дневную поверхность и т.д.

При изучении керна иногда можно наблюдать налеты и примазки нефтяных компонентов на стенках трещин. Обычно они темноокрашенные, так как представляют собой остаточные, окисленные компоненты мигрировавших через породу нефтяных флюидов: асфальтеновых и смолисто-асфальтеновых фракций. Легкие и средние компоненты (бес-

цветные и светлоокрашенные) даже при интенсивном нефтяном запахе породы остаются невидимыми.

Нефтесодержащие породы узнаются или сразу по цвету и запаху, если они сильно пропитаны нефтью, или после проверочных испытаний. Нефть может быть распределена в породе (например, в песчанике) равномерно или, чаще, неравномерно. В этом случае необходимо изучить характер ее распределения в зависимости от состава, структуры и текстуры.

Неравномерные признаки нефтенасыщения в виде «пятнистости» по всему интервалу керна чаще всего наблюдаются в переходных зонах, ближе к водонефтяным контактам или в неоднородном пласте-коллекторе с резкой изменчивостью ёмкостно-фильтрационных свойств. В этом случае необходимо детально изучить весь интервал керна на нефтенасыщенность. [4]

Проблема

Информация о керне описывается послойно: один слой - один тип породы.

В какой-то момент времени геологи поняли, что стоит детализировать описание пород: нарезали слои на фрагменты по изображениям с шагом до 1 метра и сделали для таких изображений экспертную разметку (разметка делалась несколькими экспертами, мнение которых могло не совпадать друг с другом).

Недавно геологи задумались о том, что можно размечать изображения керна с большей точностью, что позволит создавать более точные модели пластов.

1. Цели работы

Целью данной работы является получение описания керна на основе выборки фотографий. Описание выдаётся в виде текстового файла, в котором для каждого 5-ти сантиметрового участка керна указаны следующие характеристики:

- 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.

Также целью работы является написание удобной для пользователя обёртки над полученным решением для последующего использования. Это приложение будет включать в себя интерфейс, позволяющий загружать фотографии керна с описанием, на какой высоте был сделан снимок. В качестве результата будет выдана одна фотография, состоящая из склеенных фотографий керна, на которой каждый промежуток будет аннотирован своими характеристиками.

2. Постановка задачи

Для достижения приведённых целей были поставлены следующие задачи:

- Произвести разведочный анализ предоставленных данных
- Ознакомиться с возможными решениями
- Реализовать решения и найти лучшие
 - Необходимо реализовать 4 классификатора по 4-м характеристикам
- Сравнить результаты с уже имеющимися у заказчика
- Создать оболочку для удобного использования решения

3. Исходные данные

В качестве исходных данных были предоставлены фотографии кер-на и те же самые фотографии керна, но в ультрафиолетовом освещении (Рис. 3).

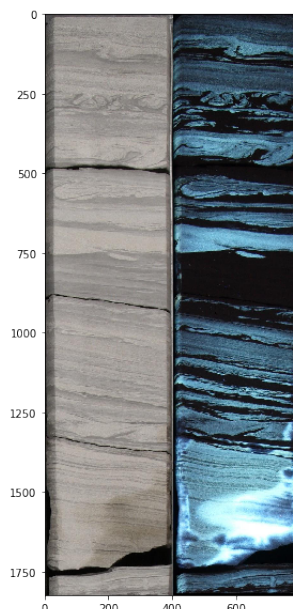


Рис. 3: Пример из исходных данных. Слева — фото керна, справа — фото того же керна, но в УФ.

К фотографиям была предоставлена таблица (Таблица 1) с описанием каждой фотографии. Нам необходимо предсказывать последние 4 параметра — **Rock** (тип породы), **Carbonate** (карбонатность), **Ruin** (разрушенность), **Saturation** (нефтенасыщенность).

Как было сообщено заказчиком, тестовая выборка будет состоять из фотографий и таблицы с двумя полями: **PhotoTop**, **PhotoDown**. Таким образом, избавимся от ненужных записей и получим таблицу (Таблица 2).

	0	1
Folder	Unload1	Unload1
Id	1000000	1000001
Field	Field6	Field6
Well	Well11	Well11
CoringTop	1957.1	1957.1
CoringDown	1963.1	1963.1
CoringTopBind	1958.3	1958.3
CoringDownBind	1964.3	1964.3
CoreRecovery	5.93	5.93
PhotoTop	0	0
PhotoDown	1	1
PhotoType	ДС	УФ
LayerTop	0	0
LayerDown	1.45	1.45
Rock	песчаник	песчаник
Carbonate	не карбонатный	не карбонатный
Ruin	не разрушен	не разрушен
Saturation	нефтенасыщенные	нефтенасыщенные

Таблица 1: Пример из таблицы исходных данных. Предоставлена информация о первых двух записях — один и тот же керн, обычная фотография и фотография в УФ.

	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 2: После удаления ненужных столбцов.

3.1. Анализ данных

Поля **PhotoUp** и **PhotoDown** означают начало и конец данной фотографии в данном образце керна. В нашем примере **PhotoUp**=0 и **PhotoDown**=1. Это значит, что длина керна на фотографии — 1 метр.

Распределение данных

В категории **Rock** очень много значений, которые имеют малое количество экземпляров в сравнении с другими категориями, поэтому выбросим их. Итоговое распределение фотографий для категории **Rock**:

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

Таблица 3: Распределение категории "тип породы".

В категории **Carbonate** всё оставим, как есть:

карбонатность	количество экземпляров
не карбонатный	4056
с карб. обломками или конкрециями	1292
слабокарбонатный	298
сильнокарбонатный	246
пятнисто карбонатный	226
среднекарбонатный	196
с примесью	100

Таблица 4: Распределение категории "карбонатность".

Категория **Ruin**:

разрушенность	количество экземпляров
частично разрушен	3132
не разрушен	2944
разрушен	338

Таблица 5: Распределение категории "разрушенность".

Категория **Saturation**:

насыщенность	количество экземпляров
не опред.	4596
нефтенасыщенные	668
пятнисто нефтенасыщенные	358
битуминозный	350
продукт	326
слабо нефтенасыщенные	116

Таблица 6: Распределение категории "насыщенность".

Всего мы имеем порядка 6000 экземпляров данных. Для решения задачи классификации этого должно быть вполне достаточно.

4. Анализ подходов

К решению задачи выделения характеристик по фотографии керна можно подойти тремя способами: классификацией изображений, распознаванием образов на изображениях и сегментацией изображений.

Для каждой из этих задач необходим алгоритм классификации.

4.1. Классификация

Задача классификации заключается в том, чтобы определить входное изображение в один из классов.

В этой задаче существуют алгоритмы машинного обучения и алгоритмы глубокого обучения. Для подходов машинного обучения необходимо сначала выделить признаки с помощью приведённых ниже алгоритмов, затем использовать техники классификации, например метод опорных векторов. С другой стороны, методы глубокого обучения способны распознавать объекты без отдельного выделения признаков. Чаще всего они основываются на свёрточных нейронных сетях.

Подходы машинного обучения для выделения признаков:

- Метод Виолы — Джонса
- Масштабно-инвариантная трансформация признаков
- Гистограмма направленных градиентов

Лучшими решениями в области являются модели глубокого обучения, такие, как свёрточные нейронные сети.

4.2. Распознавание образов

Задача распознавания образов состоит в том, чтобы определить на фотографии границы объектов известных классов. По сути то же, что и классификация, только мы с помощью определённых алгоритмов сужаем область, в которой находится объект.

Подходы:

- Region Proposals
- Single Shot MultiBox Detector
- You Only Look Once (YOLO)
- Single-Shot Refinement Neural Network for Object Detection (RefineDet)

Для более точного распознавания рекомендуется подавать на вход не только фотографию с отмеченной категорией, но и область, в которой находится данный объект.

4.3. Сегментация

Сегментация изображений — это процесс присвоения таких меток каждому пикселю изображения, что пиксели с одинаковыми метками имеют общие визуальные характеристики.

Подходы:

- Region-Based Segmentation
- Edge Detection Segmentation
- Mask R-CNN

Для ускорения обучения большинство алгоритмов помимо фотографии принимают изображение с размеченными по категориям пикселями.

Изучив статью [2], посвящённую задаче сегментации, стало понятно, что для обучения потребуются данные, в которых каждый пиксель размечен под определённую категорию. Так как этот процесс ресурсозатратный, было принято решение перейти к рассмотрению следующих подходов и вернуться к сегментации в дальнейшем.

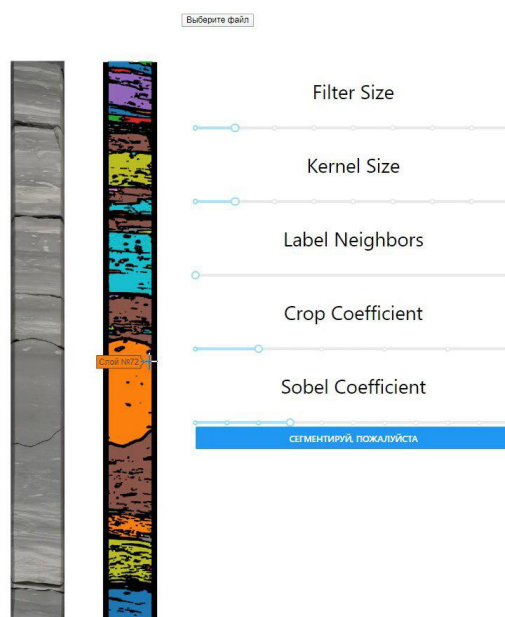


Рис. 4: Пример работы сегментационного алгоритма.

5. Решение

5.1. Подготовка данных

В каждой категории классификации количество экземпляров не сбалансировано, поэтому было принято решение применить **oversampling** (увеличение количества экземпляров в недостающих классах путём их копирования) для того, чтобы алгоритмы не учитывали перевес одних классов над другими в количестве элементов.

Категории должны быть определены с фиксированной точностью, поэтому, если 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.

5.2. Выбор алгоритмов

Попробуем решить задачу используя следующие алгоритмы:

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

5.3. 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.

На данный момент мы рассмотрим первый подход.

Протестируем и сравним точность предобученных архитектур. Для этого входные изображения преобразуем до формата входных данных обученных сетей. Результаты на примере классификации типа породы:

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

Подробнее посмотрим на ROC AUC метрику:

Топ-1 ассигасу вышла 0.75, что очень даже неплохо, учитывая, что мы используем только половину информации (модель обучена и проверена на не ультрафиолетовых фотографиях). В теории, данный подход может дать ещё большую точность, если эту модель совместить с моделью для классификации по УФ фотографиям.

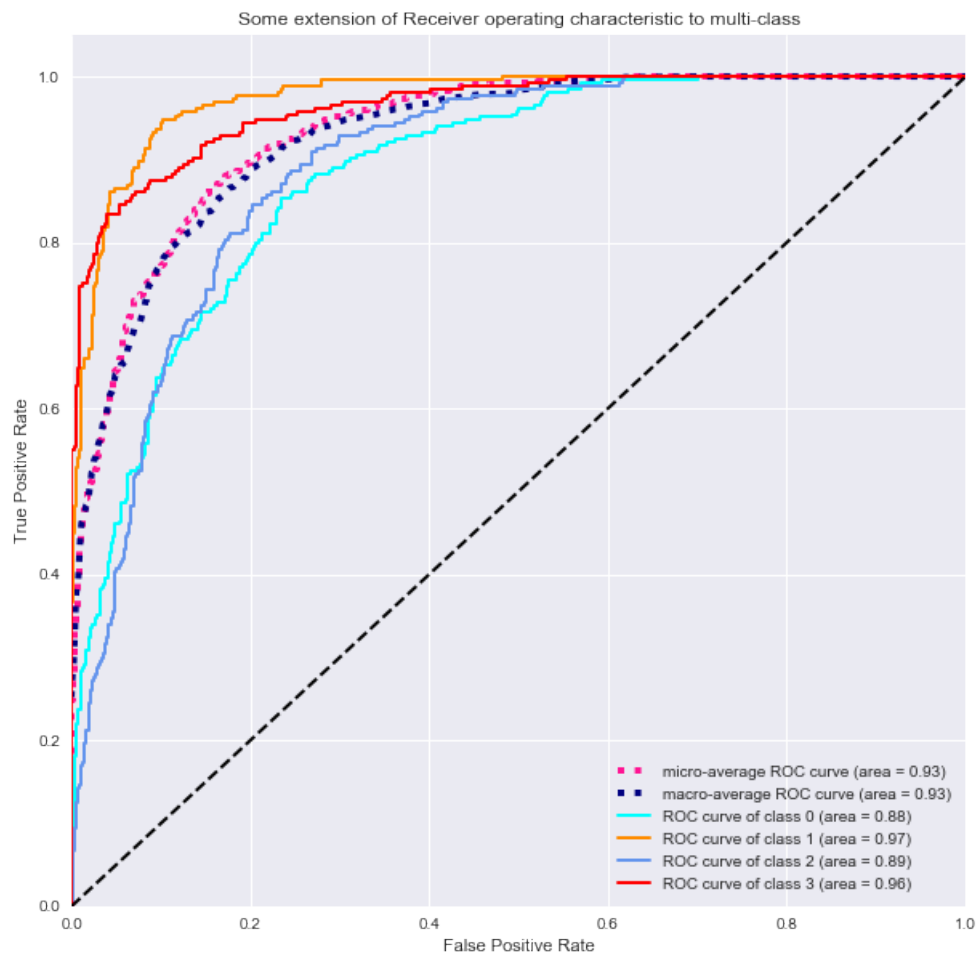


Рис. 5: VGG16 ROC AUC.

5.4. Метрическая классификация изображения

Был разобран другой подход к классификации изображений — на основе расстояния [1]. Планирую протестировать его и сравнить результаты.

Заключение

Что было сделано в рамках данной работы:

- Изучена предметная область
- Проведены изучение и подготовка данных
- Проведен разбор возможных решений
- Реализовано одно решение
 - На обычных фотографиях достигнута точность 75%
 - На ультрафиолетовых фотографиях достигнута точность 72%

Что планируется сделать:

- Улучшить реализованное решение
- Реализовать метрическую классификацию
- Реализовать сегментацию
- Сравнить результаты между собой
- Обернуть решение в удобный интерфейс

Список литературы

- [1] Gregory Koch Richard Zemel Ruslan Salakhutdinov. Siamese Neural Networks for One-shot Image Recognition // Department of Computer Science, University of Toronto. Toronto, Ontario, Canada. — 2016. — URL: <https://www.cs.cmu.edu/~rsalakhu/papers/oneshot1.pdf> (online; accessed: 19.12.2019).
- [2] Long Evan Shelhamer Jonathan, Darrell Trevor. Fully Convolutional Networks for Semantic Segmentation // IEEE. — 2016. — URL: <https://arxiv.org/pdf/1605.06211.pdf> (online; accessed: 19.12.2019).
- [3] Кузнецова Г.П. Методические приемы привязки керна к геофизическим исследованиям // ФГБОУ ВО «Российский государственный университет нефти и газа (Национальный исследовательский университет) имени И.М. Губкина» (Москва, Россия). — 2017. — URL: <http://www.neftegas.info/upload/iblock/aa8/aa814f1cd80c9dcd1301e0070462cbf5.pdf> (дата обращения: 19.12.2019).
- [4] Недоливко Н.М. Исследование керна нефтегазовых скважин // Национальный исследовательский Томский политехнический университет. — 2-е изд. — 2016. — URL: http://portal.tpu.ru:7777/SHARED/n/NEDOLIVKO/disc1/Tab2/Practicum_G12.pdf (дата обращения: 19.12.2019).