

Вопросы системной инженерии при создании информационно-аналитической системы повышения эффективности расчетов запасов углеводородов

Антон М. Тупысев
старший преподаватель
РГУ нефти и газа (НИУ) имени И.М. Губкина,
119991, г. Москва, проспект Ленинский, дом 65, корпус 1
tupysev.a@gubkin.ru

Аннотация: Предложены механизмы интеграции экспериментальных данных литологических исследований и их последующий анализ с использованием методов интеллектуального анализа данных. Описана модель базы данных образцов горных пород. Описаны аспекты системной инженерии при разработке интегрированной аналитической системы литологической информации.

Ключевые слова: литология, интеллектуальный анализ данных, оценка запасов углеводородов.

System engineering aspects in the development of information and analytical software to improve the efficiency of hydrocarbon reserves estimation

Anton Tupysev
senior lecturer
National University of Oil and Gas «Gubkin University»
65 Leninsky Prospekt, Moscow, 119991, Russia
tupysev.a@gubkin.ru

Abstract: Ways of integration of experimental data from lithological studies and their further analysis using data mining methods are considered. Data base models of core samples are described. System engineering aspects in the development of integrated information and analytical system of lithological information are described.

Keywords: lithology, data mining, hydrocarbon reserves estimation.

1 Введение

Высокий уровень неопределённости при освоении нефтегазовых месторождений связан в основном с неопределённостью в геологических данных. Несомненно, что применение современных информационных технологий для об-

работки информации является перспективным с точки зрения повышения конкурентноспособности нефтегазового производства [1].

Область литологии играет значительную роль в спектре геологических исследований. Литология – наука о составе, строении, происхождении, размещении и использовании осадочных пород [2]. В контексте нефтегазовой отрасли наиболее важными задачами данной науки представляется изучение физических, в том числе коллекторских и фильтрационных свойств пород-коллекторов, а также стратиграфические задачи.

В этом аспекте центральное место занимает характеристики порового пространства. Оценка объёмов порового пространства в совокупности с петрофизическими свойствами коллектора являются основой для освоения месторождений. За последние годы произошёл значительный скачок в исследовании порового пространства, который связан с внедрением современных инструментальных средств. К примеру, применение современных оптических микроскопов в комбинации со специализированным программным обеспечением (ПО) позволяет получать цифровые изображения шлифов и автоматизировать процесс подсчёта количественных характеристик пустотного пространства таких, как площадь, периметр и др.

В статье рассматриваются пути интеграции экспериментальных данных исследований и дальнейшего анализа с помощью средств интеллектуального анализа данных.

2 Характеристика исходных данных

Исходными данными послужили результаты исследования образцов пород. Данные были предоставлены кафедрой литологии факультета геологии и геофизики нефти и газа РГУ нефти и газа (НИУ) имени И.М. Губкина, на которой имеется значительный фонд образцов, представленных породами-коллекторами различного. Технология исследования состоит в следующем [3]. Торцевые части цилиндров, выпиленных для петрофизических исследований, подвергаются экстракции углеводородов по стандартной методике и пропитываются окрашенной смолой под вакуумом. Из полученных препаратов изготавливаются прозрачные шлифы, которые исследуются под оптическим микроскопом (пример изображения шлифа представлен на рисунке 1).

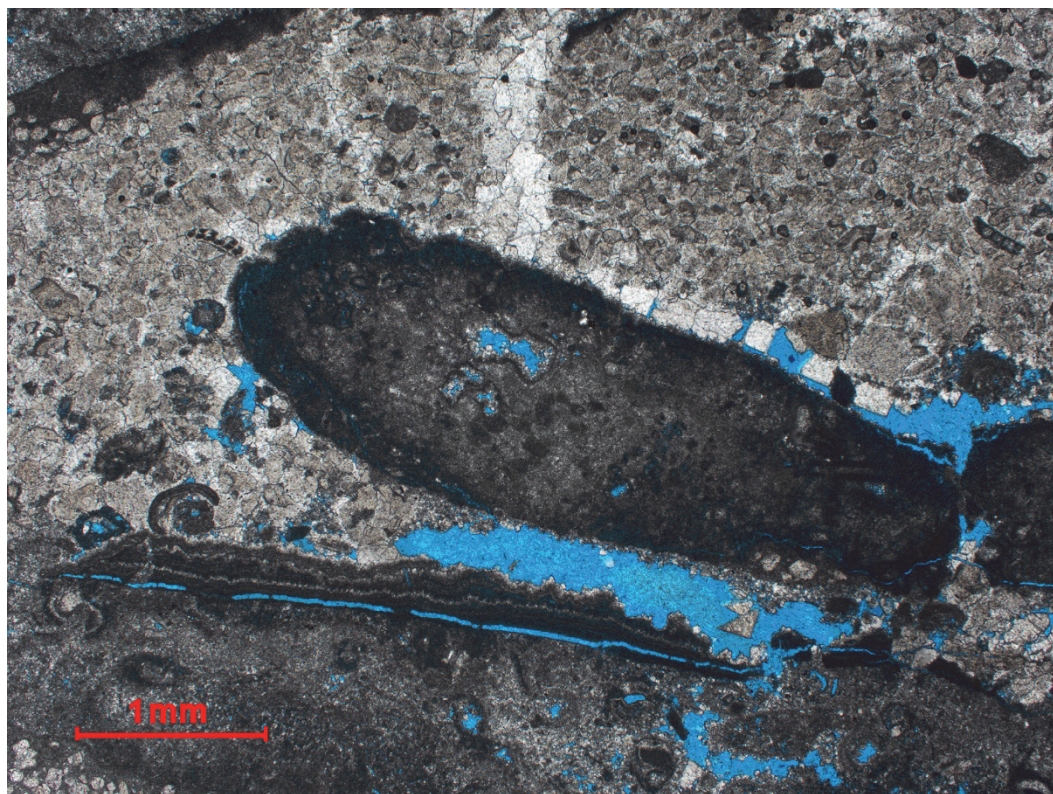


Рисунок 1 – Цифровое изображение образца породы

Цифровые изображения получены и проанализированы с помощью специализированного программного обеспечения KS-300. Помимо функций захвата изображения и его коррекции, данное ПО позволяет получить численные характеристики для каждой отдельной поры и записать данную информацию по всему изображению в электронную таблицу формата xls (MS Excel). Всего для каждой поры возможно оценить более 20 различных параметров, однако на данном этапе предполагается использовать лишь 4 – площадь поры (мкм²), длина поры (мкм), отношение ширины к дине и формфактор.

Последний параметр используется для оценки степени изрезанности границ пор. Формфактор объекта (поры или зерна) представляет собой безразмерный коэффициент $F=4\pi S/P^2$, где F – безразмерная величина формфактора, S – площадь, P – периметр объекта. Значение формфактора для круга равно 1, а для любых других форм – меньше 1.

Таким образом, был накоплен значительный массив данных в Excel-таблицах с численными характеристиками пористости изученных образцов (таблица 1), который предполагается консолидировать в единой базе данных.

Таблица 1 – Фрагмент таблицы численных характеристик пористости образца

№ поры	Глубина, м	площадь, мкм ²	длина, мкм	ширина/длина	формфактор
1	1518,66	10,36	4,74	0,55	1
2	1518,66	15,54	5,87	0,67	0,93
3	1518,66	12,09	4,74	0,83	1
4	1518,66	29,35	10,01	0,39	0,76
5	1518,66	15,54	5,58	0,71	1

3 Проектирование базы данных пористости пород-коллекторов

База данных предназначена для хранения литологической информации о пористости пород-коллекторов. По результатам проведенного анализа предметной области и исходных данных можно выделить следующие сущности, которые должны быть отражены в структуре проектируемой базы данных:

1. «FIELD» – эта сущность отводится для хранения данных о месторождении, на котором проводились те или иные исследования кернового материала;
2. «WELL». Данная сущность хранит информацию о пробуренной скважине в пределах разрабатываемого месторождения.
3. «CORE_SAMPLE». С помощью этой сущности характеризуется извлеченный из скважины керновый материал.
4. «THIN_SECTION». Рассматриваемая сущность дает описание шлифу, выпиленному из образца, полученного из керна.
5. «PHOTO» – эта сущность характеризует изображения шлифов пород, хранимые в базе данных.
6. «TYPE_PORE» – эта сущность отражает информацию о типах пор, которые присутствуют в данной породе.
7. «PORE_OLD». Одна из основных сущностей, которая содержит таблицу с количественными характеристиками пор, полученную в результате IMAGE-анализа.
8. «PORE_NEW». Рассматриваемая сущность схожа с предыдущей, единственное отличие в хранимых данных, полученных в результате подсчета параметров пор методом фрактальной размерности.

После первоначальной идентификации основных сущностей необходимо перейти к созданию концептуальной модели. Для проектирования модели базы данных была выбрана программа MicroOLAP Database Designer for PostgreSQL. Данный продукт полностью совместим с выбранной системой управления базами данных PostgreSQL, а также расширяет возможности создания базы данных.

3.1 Концептуальная модель базы данных

Концептуальная модель базы данных представляет собой описания главных (основных) сущностей, создаваемой базы, и взаимосвязей между ними [4]. На основании описанных ранее сущностей была построена следующая концептуальная модель (рисунок 2).

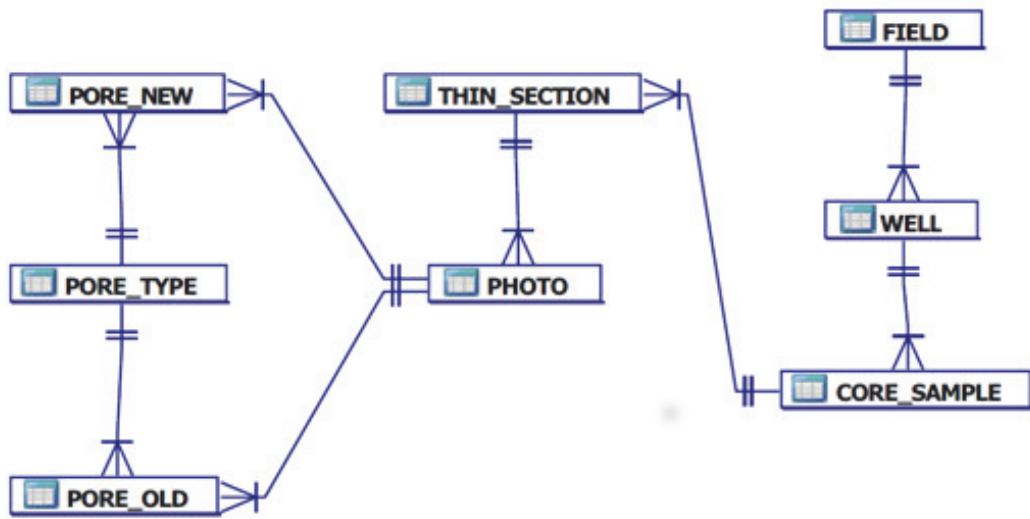


Рисунок 2 – Концептуальная модель базы данных

3.2 Логическая и физическая модель базы данных

Логическая модель расширяет концептуальную путем определения атрибутов, описания и ограничений для сущностей. Более точно идентифицируется состав сущностей и взаимосвязи между ними.

Заключительным шагом является создание физической модели данных. Физическая модель расширяет логическую путем определения для атрибутов их типов, а также формированием первичных, внешних ключей и других параметров. Полученная физическая модель представлена на рисунке 3.

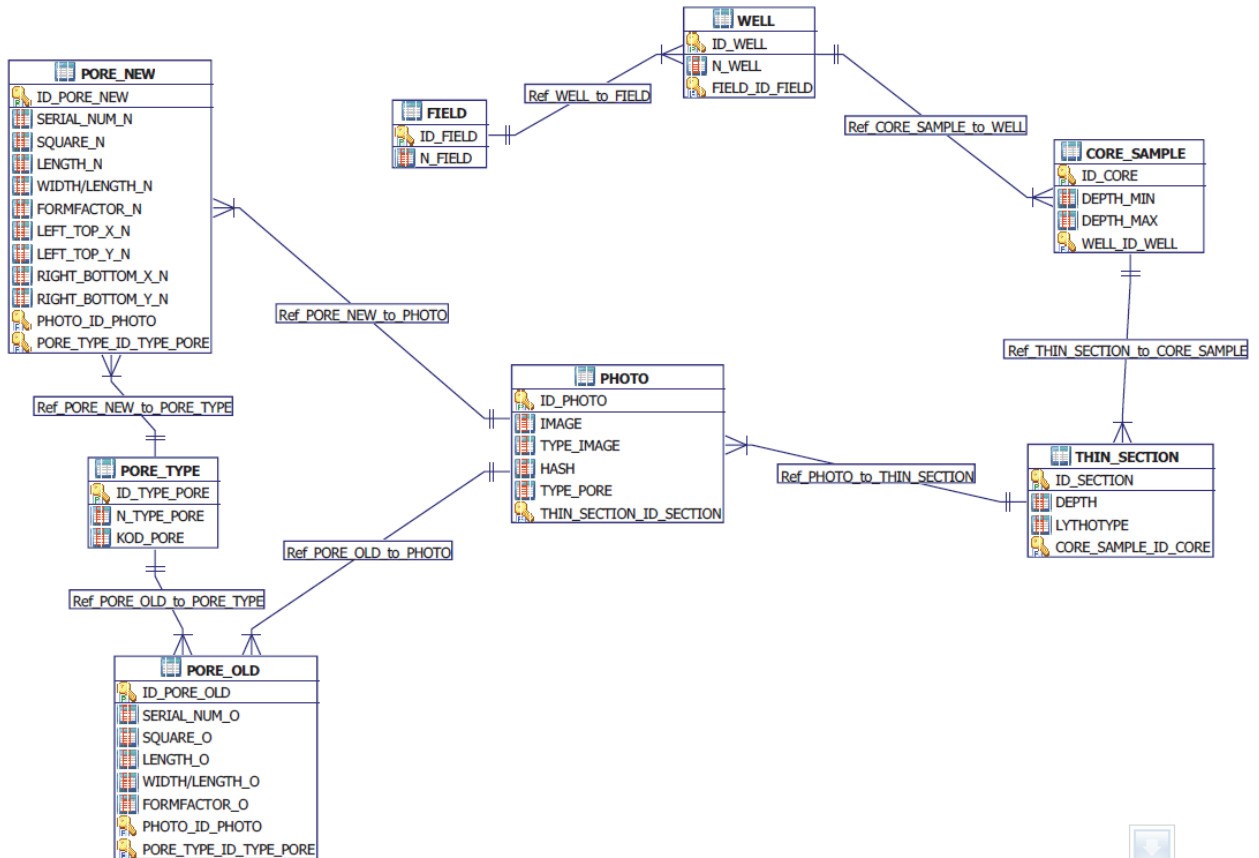


Рисунок 3 – Логическая модель базы данных

4 Фрактальный анализ цифровых изображений образцов горных пород

Высокий уровень неопределённости при освоении нефтегазовых месторождений во многом связан с разного рода ограничениями в получении и анализе геологических данных. В [5] в качестве факторов возникновения погрешностей в определении многих характеристик залежей нефти и газа и их запасов, называются недостаточная точность проводимых замеров, природная сложность изучаемых геологических объектов, ограниченность геолого-геофизической информации. В таблице 2 приводятся примеры возможных погрешностей определения параметров и запасов залежей [6].

Таблица 2 – Возможные погрешности определения параметров и запасов пластовых сводовых залежей к моменту окончания разведки

Параметры	Погрешности (\pm) определения параметров (в %) при величине балансовых запасов (в млн т)			
	до 10	10 – 25	25 – 50	50 – 300
Пористость	8,0 – 5,0	6,0 – 4,0	4,0 – 2,0	3,0 – 2,0
Запасы нефти балансовые	55,0 – 30,0	35,0 – 20,0	30,0 – 15,0	20,0 – 10,0
Запасы растворенного газа балансовые	50,0 – 30,0	37,0 – 20,0	32,0 – 15,0	22,0 – 12,0
Запасы нефти извлекаемые	65,0 – 35,0	45,0 – 25,0	35,0 – 20,0	25,0 – 13,0
Запасы растворенного газа извлекаемые	67,0 – 37,0	46,0 – 27,0	36,0 – 22,0	27,0 – 15,0

Учёт фрактальных свойств геологических систем, применение соответствующих численных методов и разработка программного обеспечения на их основе позволяет повысить качество принимаемых управленческих решений на этапе разработки месторождения.

Многие исследования пластовых систем, содержащих углеводородные полезные ископаемые, показали применимость аппарата фрактальной геометрии в анализе процессов осадконакопления, формировании трещин, оценке пористости пород, обработке временных рядов при гидродинамических исследованиях скважин [напр., 7,8]. В [9] показан самоподобный характер исследуемых геологических структур и процесса добычи флюида. В [10] показана связь фрактальной размерности пористо-пустотного пространства пород с геометрическими свойствами и величиной пористости.

В рамках данной работы теория фракталов применяется для корректировки расчётов линейных характеристик порового пространства на образцах горных пород. разработанный модуль был опробован для расчёта фрактальной размерности пор на цифровых изображениях образцов горных пород. В исследовании использовались фотографии образцов породы, преобразованные в бинарные чёрно-белые изображения, очищенные от любых объектов кроме пор (рисунок 4).



Рисунок 4 – Цифровое изображение образца горной породы (чёрные объекты соответствуют порам)

Для 268 изображений образцов горных пород вычислены фрактальные размерности по методу триангуляции и методу box-counting. Результаты, полученные методом триангуляции, показали расхождения с методом box-

counting в пределах 6%, что подтверждает высокую чувствительность первого метода к выбору начальных параметров. Помимо значения шага точность зависит от выбора начальной и конечной длины разбиения. Рекомендуется брать следующие параметры, при которых получились надёжные результаты: интервал разбиений [40; 100] и шаг $h=1$.

5 Интегрированная система хранения и обработки литологической информации

Описанная база данных является основой комплексной системы хранения и обработки литологической информации. В состав системы базы исходных данных, комплекс методов Data Mining и интерфейс для работы с ними.

Комплекс методов Data Mining подразумевает использование ряда алгоритмов (корреляционный анализ, кластерный анализ и др.). Интерфейс должен быть специально разработан с целью удобного и простого применения этих методов к данным из базы, а также для наглядного представления результатов анализа.

На рисунке 5 представлена принципиальная блок-схема предлагаемой системы.

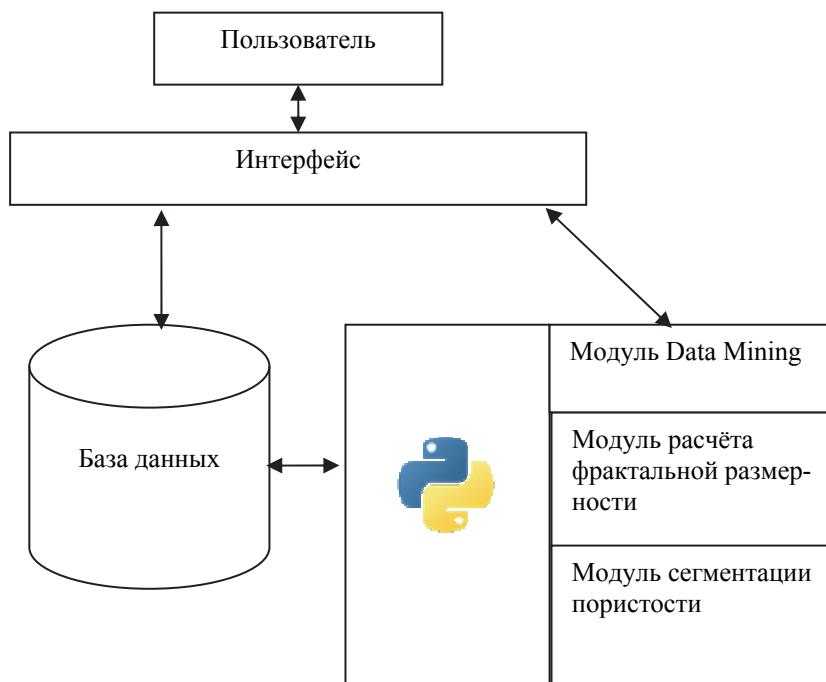


Рисунок 5 – Блок-схема предлагаемой системы

6 Сегментация пустот на цифровых изображениях

В состав аналитического модуля предлагается включить в разрабатываемую систему модуль автоматизированной сегментации внутрiformенных и межформенных пустот по изображению шлифа породы коллектора. Разработанный алгоритм сегментации пустот, представляющий из себя модификацию архитектуры Fast-SCNN, способен целиком обрабатывать цифровые изображения петрографических шлифов.

Изображения были приведены к одному межпиксельному расстоянию в 4 мкм, что в данном случае соответствует разрешению одного кадра в 3640x4019 пикселей. Тестовая выборка была составлена путем случайного выбора 20 снимков.

Изображения из обучающей выборки были разрезаны на образцы размером 512x512 пикселей с шагом 384 пикселя по осям X и Y. Итоговое количество образцов в обучающей выборке составило 4176 шт. Изображения из тестовой выборки использовались для оценки лучшей эпохи, они были разрезаны на образцы размером 1024x1024 пикселей с шагом 512 пикселя по осям X и Y.

Аугментация обучающей выборки производилась путем случайных отражений, случайных поворотов и упругой деформации [11]. Значения пикселей каждого изображения были приведены к интервалу $[-1; 1]$ путем масштабирования.

Для обучения модели, построенной по предложенной архитектуре, были использованы следующие параметры:

- алгоритм оптимизации Adam [12] со скоростью обучения 0,001;
- batch Size равен 24;
- создание контрольной точки на эпохе с лучшим средним коэффициентом Дайса.

Для обучения моделей использовалась ЭВМ со следующими техническими характеристиками в таблице 3.

Таблица 3 – Технические характеристики ЭВМ

Параметр	Характеристика
CPU	2x Intel Xeon E5-2620 v4
GPU	2x NVIDIA GeForce GTX 1080 Ti (11GiB)
RAM	256 GiB

На рисунке 6 представлены два примера результатов работы алгоритма (столбцы изображений слева и справа) в сравнении с другими подходами. Здесь верхние изображения – исходные образцы, затем сверху вниз: экспертная разметка "вручную", разметка UNet, разметка Fast-SCNN, разметка модифицированной Fast-SCNN.

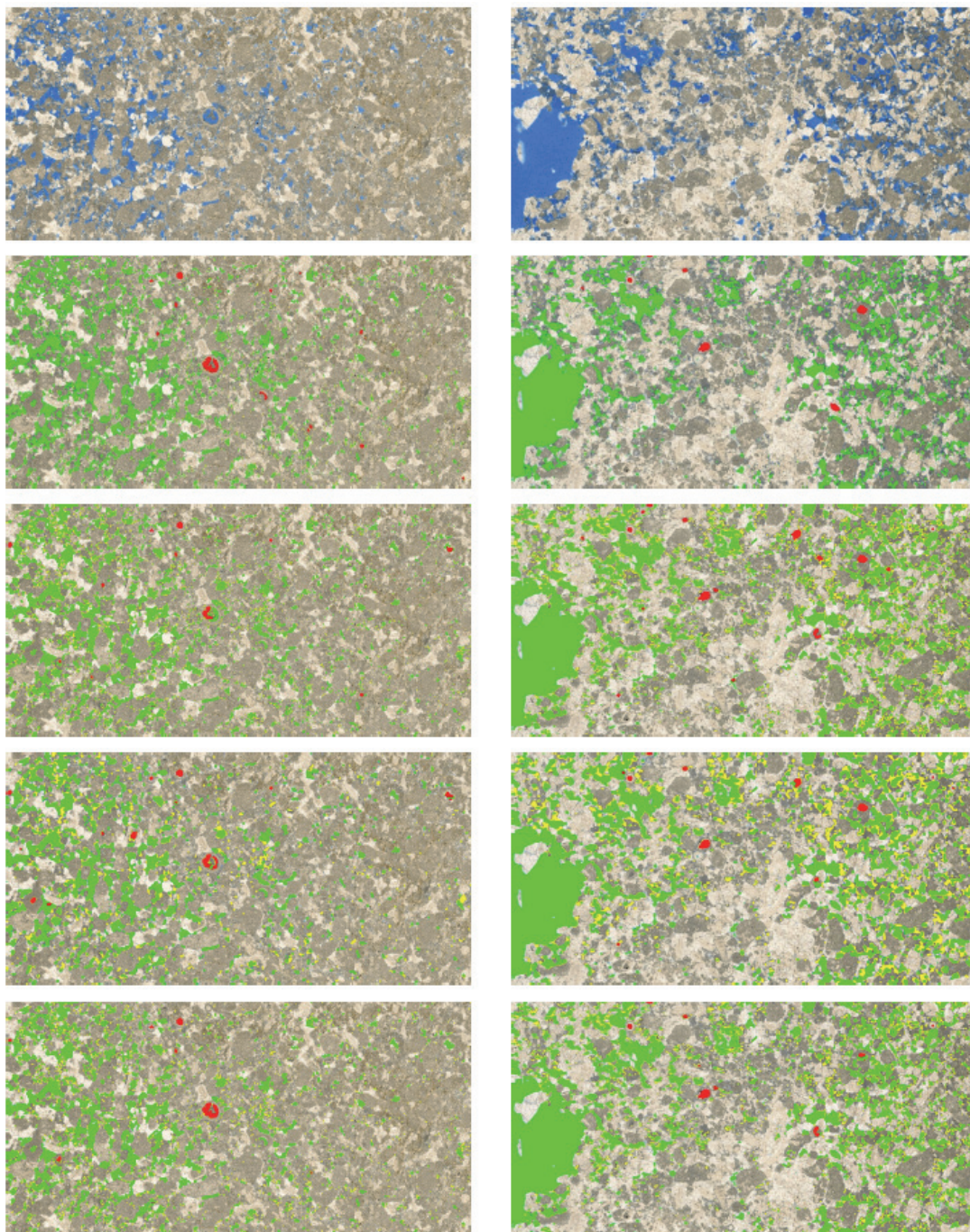


Рисунок 6 – Примеры результатов сегментации

Были получены следующие коэффициенты Дайса для трех архитектур, представленные в таблице 4.

Таблица 4 – Коэффициенты Дайса для тестовой выборки

Метод	Коэффициенты Дайса (безразмерные величины)				Среднее арифметическое параметров
	Тип поро-ды	Пустоты	Микропористость	Внутриформенные пустоты	
UNet	0,9735	0,8322	0,322	0,4237	0,6382
Fast-SCNN	0,9490	0,7095	0,1773	0,3709	0,5517
Предложенный метод	0,9710	0,8028	0,2588	0,4767	0,6273

Согласно полученным результатам, архитектура UNet показала лучший результат в сегментации микропористости, при этом предложенная модель выиграла в сегментации внутриформенных пустот. Дополнительно была проведена оценка размера модели для каждого метода, а также количество операций и затраченное время (см. таблицу 5) на обработку одного RGB изображения разрешением 1024x1024px.

Таблица 5 – Особенности работы методов

Метод	Количество параметров (млн. шт.)	Количество операций вывода (FLOP)	Время вывода для CPU (сек)	Время вывода для GPU (сек)
UNet	7,7761	4329,734G	2.9265	0.0411
Fast-SCNN	1,1361	4.004G	0.2291	0.0074
Предложенный метод	1,1351	13.420G	1.2330	0.0305

Максимальная пространственная детализация сегментационной карты достигнута алгоритмом UNet, который также хорошо сегментирует малую внутрифирменную пористость, но при этом имеет проблемы с сегментацией пустот больших скелетных форм. Архитектура Fast-SCNN имеет наименьшую детализацию. Предложенная архитектура решает проблему детализации Fast-SCNN, хорошо сегментирует крупную внутриформенную пористость, и в целом показывает хорошие результаты на остальных классах пустот.

Несмотря на хорошие показатели, у модели UNet высокая вычислительная сложность. Предлагаемая архитектура имеет на порядок меньшее количество операций и менее требовательна по памяти, при сравнимом качестве разметки. Обработка полноразмерного снимка алгоритмом UNet на GPU не оказалась возможной, поскольку модель требует более 11GiB доступной видеопамати.

7 Заключение

Для нефтегазовой отрасли, играющей существенную роль в российской экономике, объем и разнообразие геологических данных служат отправной точкой для развития конкретных технологий: бурения и разработки углеводородных месторождений; перерабатывающей промышленности и нефтегазохимических производств и др.

Нефтегазовые месторождения формируются в течение сотен миллионов лет. Это эволюционные и трудно формализуемые процессы, поэтому на протяжении многих лет геология рассматривалась как описательная наука.

С середины прошлого века с появлением и развитием информационных технологий появилась возможность получения, сбора, хранения обработки разнородных данных, математического описания ряда процессов. Появились средства и возможности математизации сложных эволюционных процессов, геологических объектов.

В этих условиях становится актуальным привлечение методов системного анализа к обработке геологических данных, использование принципов и стандартов системной инженерии (Национальный стандарт Российской Федерации. ГОСТ Р 15793-2016 Системная и программная инженерия, Процессы жизненного цикла).

В статье на основе достижений в области системного анализа и современных ИТ предлагается архитектура информационно-аналитической системы, обеспечивающей переход от хранения данных в отдельных электронных таблицах к комплексному решению. Предусмотрена возможность обработки полученных данных методами интеллектуального анализа, что может помочь в дальнейшем перейти к многомерному анализу, т.е. к качественно новой оценке имеющихся экспериментальных данных. Этот этап развития информационных технологий представляется закономерным шагом в автоматизации проводимых литологических исследований.

Литература

- [1] Григорьев Л.И., Кершенбаум В.Я., Костогрызов А.И. Системные основы управления конкурентоспособностью в нефтегазовом комплексе. – М.:Изд-во НИИНГ. 2010
- [2] Фролов В.Т. Литология. Кн.1: Учебное пособие. – М.: Изд-во МГУ, 1992.
- [3] Постников А.В. Комплексное изучение структуры порового пространства органогенных известняков/ А. В. Постников [и др.] // Инновационное развитие нефтяной и газовой промышленности России: наука и образование: губкинские чтения (24-25 ноября 2009 г., г. Москва), 18-е. РГУ нефти и газа им. И.М. Губкина. – М., 2009
- [4] Коннолли Т., Бегг К. Базы данных. «Проектирование, реализация и сопровождение. Теория и практика» — 3-е издание.: Пер. с англ. – М.: «Вильямс», 2003.
- [5] Бакиров А.А., Бакиров Э.А., Габриэлянц Г.А., Керимов В.Ю., Мстиславская Л.П. Теоретические основы поисков и разведки нефти и газа. Кн 2: Методика поисков и разведки скоплений нефти и газа – М.: ООО «Издательский дом Недра», 2012.
- [6] Справочник по нефтепромысловой геологии [под ред. Быкова Н.Е., Максимова М.И., Фурсова А.Я.] – М.: Недра, 1981.
- [7] Flake Garry William, The Computational Beauty of Nature. Computer Explorations of Fractals, Chaos, Complex Systems, and Adaptation. – The MIT Press, Cambridge, Massachusetts, London England, 2001.
- [8] Митюшин А.И., Разбегина Е.Г., Фрактально-статистический анализ процесса добычи углеводородов, -- М., МАКС-Пресс, 2010.
- [9] Григорьев Л.И., Мухина А.Г., Изюмов Б.Д. Формирование модели «Жизнь пласта» для повышения эффективности управления разработкой месторождения // ВЕСТНИК ЦКР РОСНЕДРА. – 2015. – №4. – С. 6-15.
- [10] Кадыров Р.И., Закиров Т.Р. 2D фрактальный и мультифрактальный анализ пустотно-пористого пространства карбонатных коллекторов нефти // НЕФТЯНОЕ ХОЗЯЙСТВО. – 2016. -- №11. – С. 72-74.
- [11] Simard P.Y., Steinkraus D., Platt J.C. Best practices for convolutional neural networks applied to visual document analysis // Icdar. 2003. Vol. 3.
- [12] Kingma D.P., Ba J. Adam: A method for stochastic optimization // arXiv preprint arXiv:1412.6980, 2014.