

Литологическая интерпретация фотографий керна на основе компьютерного анализа

А. В. Ивченко

Московский физико-технический институт
ivchenko.a.v@phystech.edu

Е. Е. Барабошкин¹, Л. С. Исмаилова²,

Д. М. Орлов³, Д. А. Коротеев⁴

Сколковский институт науки и технологий

¹Evgenii.Baraboshkin@skoltech.ru,

²L.Ismailova@skoltech.ru, ³D.Orlov@skoltech.ru,

⁴d.koroteev@skoltech.ru

Е. Ю. Барабошкин

Московский государственный университет имени М.В. Ломоносова
ejbaraboshkin@mail.ru

Аннотация. Работа посвящена анализу изображений срезов горных пород с помощью методов машинного обучения и нейронных сетей. Возрастающий объем информации и требования к скорости обработки данных, отсутствие отраслевых и государственных стандартов, а также тяжело различимые особенности и субъективизм в геологических задачах делает актуальным создание системы автоматического унифицированного анализа керна материала. В статье описывается автоматизация анализа геологической информации на основе анализа изображений горных пород с помощью методов машинного обучения и нейронных сетей. Демонстрируется способность алгоритма различать шесть литотипов с максимальной точностью 98%.

Ключевые слова: машинное обучение; нейронные сети; седиментология; керн, фотографии; описание; сегментация; литология; автоматизация

1. ВВЕДЕНИЕ

Объем данных о строении недр возрастает. Данные включают в себя геофизические, геохимические, геологические и другие методы исследования. Многие из этих направлений обрастают всё новыми данными, что связано с непрерывным развитием технологий разведки месторождений. Требования к скорости обработки информации также возрастают. Нефтяными компаниями ставится вопрос о максимальной оптимизации процесса разведки-добычи без расширения штата компании, занимающегося геологоразведочными работами (ГРП) [9]–[11], [18], [18]–[21]. Этот вопрос может быть решен путем автоматизации различных процессов при помощи алгоритмов машинного обучения. Подобное решение позволяет избавиться от субъективной оценки при описании различных параметров и создать единую, унифицированную схему для описания получаемых в результате исследований данных. Наша работа нацелена на автоматизированный анализ керна.

А. Литературный обзор

Существует несколько работ, посвященных автоматизации анализа керна материала. В 2012 году была опубликована работа А.К. Пателя, в которой проводилась классификация пород с помощью методов SVR [16]. Выделялось 6 типов породы с ошибочными результатами в 6–10%. Недостатком работы являются унифицированные входные данные, собранные одним исследователем в едином виде, а также отсутствие анализа работы на реальных геологических образцах со сложными структурными и текстурными особенностями.

И.С. Хасанов выпустил ряд работ [2]–[7], посвященных полуавтоматической обработке фотографий керна горных пород. Изображение преобразуется в матрицы в различных цветовых схемах (RGB или HSL), после чего каждому классу цвета оператором присваивается та или иная характеристика породы (пористость, проницаемость и т.д.), присущая этому цвету, и ее литотип. Далее исследуются вариации параметров отдельных цветов и гистограммы распределения этих цветов. Исследование предусматривает обработку изображений, полученных как при дневном свете, так и при ультрафиолетовом. Недостатком работы является необходимость активного привлечения к работе специалиста-геолога с программными инструментами, только лишь ускоряющими его работу.

Близкой к работе Хасанова является работа В.Е. Элингтона «Products and methods for identifying rock samples» [14]. Цветность и тип породы анализируются таким же образом. Для каждого отдельного образца породы определяются его цветовые характеристики. Далее анализируется вся скважина с составлением размеченной по цвету литологической колонки. Выделяются граничные значения цветов, соответствующие разным литотипам.

Наиболее близкой к данной работе является исследование 2017 года, проведенное М.М. Межгани [15].

Анализируются образцы керна, сфотографированные при дневном свете в высоком разрешении, производится анализ данных, полученных от пластового микроимиджера (FMI) с применением компьютерного анализа. Недостающие части керна на фотографиях (к примеру, изъятые образцы в виде цилиндрических проб, полная терминология см. [1]) заполняются согласно окружающим текстуре и структуре по методу многоточечной статистики (multi-point statistics – MPS), что требуется для алгоритма анализа, но уменьшает точность. Основным недостатком работы является проведения сравнения с библиотекой образцов путём выделения в явном виде таких параметров, как контрастность, интенсивность пикселей, что не является объективным описанием горных пород, а также зависящих от условий съёмки. Отсутствует описание изображений, сделанных при ультрафиолетовом освещении.

В. Цели и задачи

Целью исследования является создание алгоритма способного за непродолжительное время (минуты, часы) произвести определение типа породы в автоматическом режиме, что необходимо в условиях обработки большого количества данных.

В задачи исследования входят: определение перечня изучаемых параметров керна, их классификация и группировка; сбор материала для обучения алгоритма; создание обучающей выборки; создание и настройка алгоритма.

II. МАТЕРИАЛЫ И МЕТОДИКА

А. Подготовка данных

Для упрощения обработки данных было создано дерево признаков литотипов. Это повысило точность анализа и уменьшило вычислительную сложность.

Существенной частью проекта является предварительная обработка изображений керна с целью подготовки для анализа нейронной сетью, включающая в себя сегментацию изображения на отдельные фрагменты согласно заданному шагу, предварительное извлечение изображений из стандартных форматов хранения геологических отчетов XLS и Corel (рис. 1 и рис. 3).

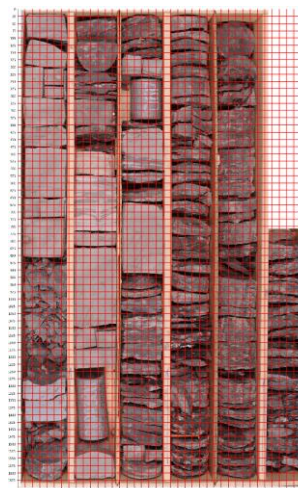


Рис. 1. Фрагмент сегментированных данных

В. Работа анализатора

Анализатор принимает на вход цветное или в градациях серого изображение и преобразует в n-мерную матрицу распределения численных значений цвета и нормализует эти значения в интервале от 0 до 1. Дальнейший анализ включает в себя сегментацию, уменьшение размерности матрицы (свертки) и другие преобразования.

В конечном итоге глубокая нейронная сеть сегментирует всё изображение на отдельные фрагменты согласно заданному шагу. Каждому сегменту присваивается вероятность принадлежности к тому или иному типу породы или иной определяемой характеристике (пористость, проницаемость и т.д.). Исходя из этого исследуемому образцу присваивается наиболее вероятная конкретная характеристика. На выходе производится визуализация результатов в виде изображения с размеченной картой определяемых свойств (рис. 2).

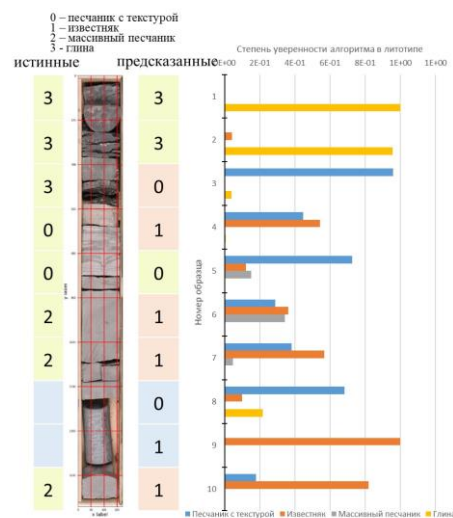


Рис. 2. Пример работы анализатора

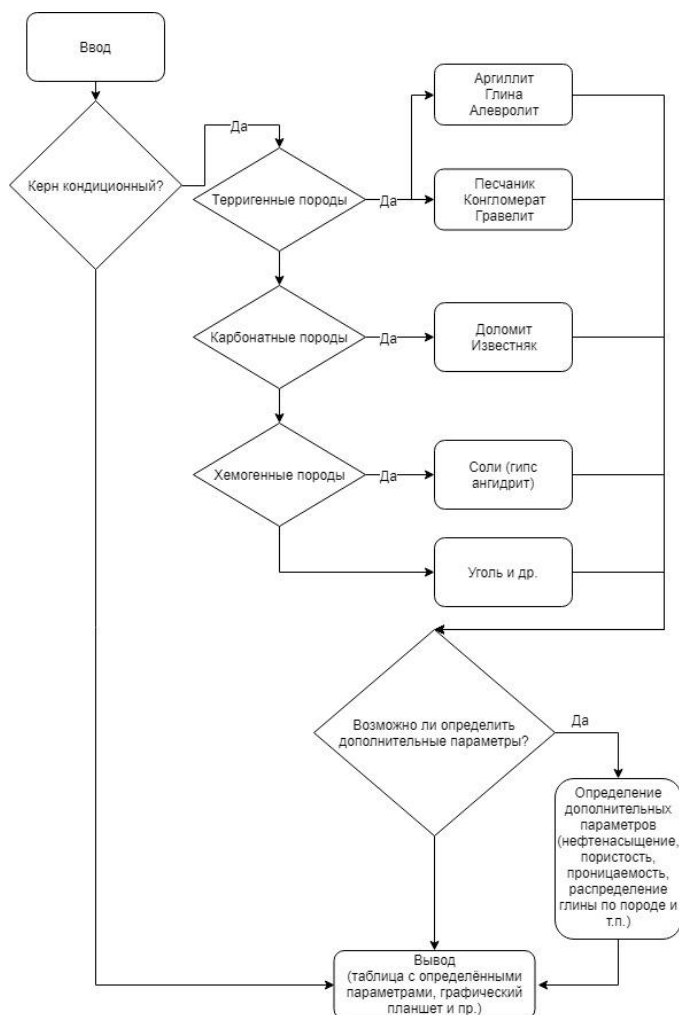


Рис. 3. Принципиальная схема работы отдельного блока системы.

Детальность карты зависит от задачи исследования и задаётся исследователем. В дополнение к этому в выводе могут присутствовать таблицы данных и данные в других форматах с определёнными свойствами.

В качестве ядра была выбрана открытая программная библиотека для машинного обучения Google TensorFlow [23]. Ряд экспериментов с оболочками для ядра привел к выбору библиотеки Keras [8]. Предварительная сегментация изображений производится на языке Python с помощью библиотеки PIL (рис. 1) [17]. Обработка массивов данных ведётся с помощью библиотеки numpy [12], ряд алгоритмов машинного обучения взята из библиотеки scikit-learn [19]. Работа непосредственно с изображениями ведётся с помощью библиотеки OpenCV [13]. Для повышения точности работы нейронной сети и уменьшения риска переобучения применен алгоритм Dropout [22].

На их основе была написана глубокая сверточная нейронная сеть [10]. Производились различные эксперименты с глубиной сети, количеством нейронов на слоях, функциями распространения ошибок, размеров изображений. Сеть показывает 84% точности работы в пределах одного класса при размере изображений 128x128



Рис. 4. Пример исходных данных слева направо: слоистый песчаник, известняк, разрушенный керн, массивный песчаник, нераспиленный керн, аргиллит.

пикселей при равномерном распределении количества образцов. Для обучения и проверки результатов работы нейронной сети использовалась собранная из различных источников выборка данных, состоящая из шести литотипов (рис. 4):

- слоистого песчаника;
- известняка;
- разрушенного керна;
- не разрезанного керна;
- массивного песчаника;
- аргиллита.

Фотографии были сделаны при дневном и ультрафиолетовом свете в различных условиях.

III. РЕЗУЛЬТАТЫ

Для обучения нейронной сети было использовано 800 изображений, приведённых к размеру 512x512 пикселей, изначально составляющих 6 классов. В дальнейшем количество классов было уменьшено до 4, так как класс разрушенных пород и не разрезанного керна определялись верно с точки зрения геологии (тип породы верен) и не верно с точки зрения машины (класс разрушенный керн не верно определён как порода). Количество материалов по классам неравномерно: наибольшее количество материала по аргиллиту (335 образцов), наименьшее – по массивному песчанику (43 образца). В качестве обучающей выборки было использовано 60% собранного материала и, соответственно, 40% составляла тестовая выборка. Результаты показывают до 96% точности и 98% полноты (рис. 5). Однако на некоторых классах (известняк, 124 образца) результаты падают до 42% точности.

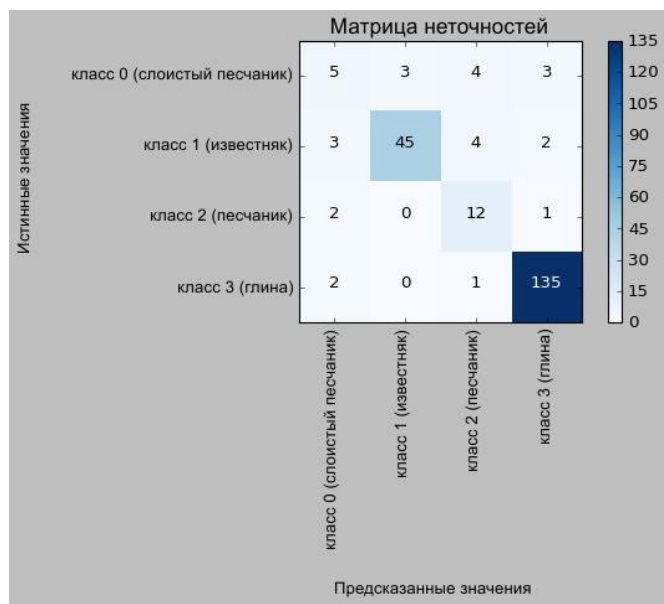


Рис. 5. Матрица неточностей для 4 классов

ВЫВОДЫ

Полученные результаты демонстрируют возможность применения предложенного подхода автоматического анализа изображений керна с помощью свёрточных нейронных сетей. Дальнейшие работы будут направлены на увеличение базы геологических данных для обучения сети, оптимизацию работы сети с учётом роста количества изображений, корректную аугментацию данных. Дополнительные работы будут вестись с обработкой литотипов, на которых система показывает низкие результаты.

Также будут добавлены модули поиска и анализа объектов на керне, в частности – анализ размер зерна, привязка данных к глубине отбора керна.

ВЫРАЖЕНИЕ ПРИЗНАТЕЛЬНОСТИ

Выражаем благодарность Елене Анатольевне Жуковской за консультацию, Марселию Харитонову Хайруллину за предоставленные материалы.

СПИСОК ЛИТЕРАТУРЫ

- [1] Андерсен М.А., Дункан Б., МакЛин Р. Анализ керна: истина в последней инстанции при оценке характеристики пласта // Нефтегазовое обозрение. 2013. 25, № 4. С. 4–17.
- [2] Постников А.В., Хасанов И.И. Компьютерное моделирование структурно-текстурных особенностей пород // Материалы Балтийской школы-семинара «Петрофизическое моделирование осадочных пород, Петромодель 2015». Санкт-Петербург. 14 – 18 сентября. 2015.
- [3] Постников А.В., Постникова О.В., Оленова К.Ю., Сивальнева О.В., Хасанов И.И., Осинцева Н.А., Ганаева М.Р. Новые методические аспекты литологических исследований пород баженовской свиты // Нефтяное хозяйство. 2015. №10. С. 23–27.

- [4] Хасанов И.И. Анализ цветности пород по цифровым изображениям керна // Геология нефти и газа. 2014. № 5. С. 33–39.
- [5] Хасанов И.И. Методика количественной оценки цветности кернового материала и её практическое применение. // Материалы Балтийской школы-семинара «Петрофизическое моделирование осадочных пород, Петромодель 2013» 2013.
- [6] Хасанов И.И. Применение компьютерного анализа цифровых изображений для исследования кернового материала (тезисы) // 17-я конференция по вопросам геологоразведки и разработки месторождений нефти и газа EAGE «Геомодель-2015». Геленджик. 7-10 сентября. 2015.
- [7] Хасанов И.И., Пономарев И.А., Постников А.В., Осинцева Н.А. Методика количественной оценки емкостных параметров пород-коллекторов с применением цифровой обработки фотографий керна (тезисы) // 18-я конференция по вопросам геологоразведки и разработки месторождений нефти и газа EAGE «Геомодель-2016». Геленджик. 12-15 сентября. 2016.
- [8] Chollet F., Google, Microsoft, others. Keras. // GitHub repository. 2015. URL: <https://keras.io>
- [9] Gurevich I.B., Yashina V.V. Descriptive Image Analysis: Genesis and Current Trends. // Pattern Recognit. Image Anal. 2017. № 27. P. 653–674.
- [10] Lecun Y., Boser B., Denker J., Henderson D., Howard R.E., Hubbard W., Jackel, L. Handwritten Digit Recognition with a Back-Propagation Network. // Neural Information Processing Systems. 1997. №2.
- [11] Lu B., Cui M., Liu Q., Wang Y. Automated grain boundary detection using the level set method. // Computers & Geosciences. 2009. №35(2). 267-275.
- [12] NumPy. Package for scientific computing with Python. // Version 1.14.2. URL: <http://numpy.org/>
- [13] OpenCV. Open Source Computer Vision Library // Version 3.4.0. URL: <https://opencv.org/>
- [14] Pat. US 2013/0156270 A1. / W.E.J.Ellington, J.C.Moore, M.A. Smith, Dubinsky G.L. Products and methods for identifying rock samples. 2013.
- [15] Pat. US2017286802 / M.M. Mezghani, S.H. Shammari, F.A. Anifowose. Automated core description. 2017.
- [16] Patel A.K., Chatterjee S., Gorai A.K. Development of online machine vision system using support vector regression (SVR) algorithm for grade prediction of iron ores //2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA). May 8-12. 2017. / IEEE, 2017. P. 149–152.
- [17] Pillow. Python Imaging Library. // Version 5.0.0. URL: <http://pillow.readthedocs.io/>
- [18] Prince C.M., Chitale J. Core Image Analysis: Reliable Pay Estimation in Thin-Bedded Reservoir Units. // Soc. of Core Analysts Annual Meeting. 2008. P. 1-6.
- [19] Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J., Passos A., Cournapeau D., Brucher M., Perrot M., Duchesnay E. Scikit-learn: Machine Learning in Python // JMLR 12. 2011. P. 2825–2830. Version 0.19.1. <http://scikit-learn.org>
- [20] Rabbani A., Ayatollahi S. Comparing three image processing algorithms to estimate the grain-size distribution of porous rocks from binary 2D images and sensitivity analysis of the grain overlapping degree. // Special Topics and Reviews in Porous Media. 2015. №6. P. 71-89.
- [21] Richa R., Mukerji T., Mavko G., Keehm Y. Image analysis and pattern recognition for porosity estimation from thin sections. // Seg Technical Program Expanded Abstracts. 2006. P.1968-1972.
- [22] Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdinov R.. Dropout: A Simple Way to Prevent Neural Networks from Overfitting // Journal of Machine Learning Research. 2014. №15. P. 1929-1958.
- [23] TensorFlow. An open source machine learning framework for everyone. // Google. Version 1.6.0. URL: <https://tensorflow.org/>