

ПРИМЕНЕНИЕ СРЕДСТВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ИССЛЕДОВАНИЯ ОБРАЗЦОВ КЕРНА

Денисов В.И.¹, Задорожных Е.А.²

¹ *Томский политехнический университет, ИШИТР, зр. А2-39, vid3@tpu.ru*

² *Томский политехнический университет, ИЯТШ, зр. 0B02, eaz28@tpu.ru*

Аннотация

В докладе проводится анализ работ по автоматизации исследования керна на основе обработки и интерпретации его изображений. Полученные данные позволяют разработать наиболее эффективный подхода к анализу изображений керна и на его основе решить задачу сегментации образцов керна на дневных изображениях.

Ключевые слова: сверточная нейронная сеть, U-Net, керн, автоматизация, машинное обучение.

Введение

Керн представляет из-за себя некоторый объем горной породы, обычно в виде образца цилиндрической формы. Именно на основе анализа керна устанавливаются значения характеристик пласта и выводятся зависимости, позволяющие строить гидродинамические модели месторождений, без которых сейчас немислим сам процесс разработки недр [1].

При этом, сама работа по исследованию образцов керна, сопряжена с большим количеством монотонных однотипных операций. Поэтому автоматизация процесса исследования образцов керна на сегодняшний день является одной из важнейших задач цифровизации керна.

В общем случае процесс автоматизации можно представить в виде последовательности следующих этапов:

- анализ существующих подходов к исследованию образцов керна и формирование на их основе наиболее эффективного подхода, а также требований к нему;
- сегментация образцов керна на дневных изображениях;
- сегментация образцов керна на ультрафиолетовых изображениях;
- сегментация ультрафиолетового свечения;
- определение нефтенасыщенных интервалов на изображения керна с использованием дневных и ультрафиолетовых фотографий.

Анализ методов исследования изображений керна

К настоящему времени существует большое количество работ, в которых в той или мере рассмотрены методы автоматизации исследований образцов керна. Остановимся более подробно на основных работах с целью определения преимуществ и недостатков различных подходов.

1. Применение вероятностной нейронной сети (PNN)

Данный подход к исследованиям изображений керна был применен А.К. Пателем в работе «Computer vision-based limestone rock-type classification using probabilistic neural network» [2]. В статье описывается процесс автоматизации описания, анализа и классификации типов известняковых пород на основе вероятностной нейронной сети. Ценность рассматриваемой работы заключается в том, что автор отходит от наиболее часто используемой в подобных работах сверточной нейронной сети и применяет PNN. Сама архитектура предложенной автором методологии изображена на рис. 1.

Стоит отметить, что автором в данной работе использовались унифицированные входные данные, собранные одним исследователем в едином виде. Кроме того, отсутствует анализ применения нейронной сети на реальных геологических образцах со сложными структурными и текстурными особенностями.

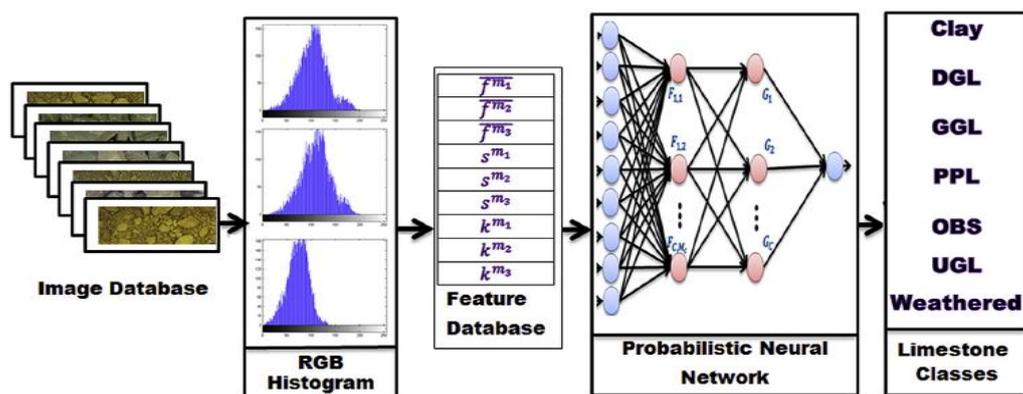


Рис. 1. Методология системы в работе «Computer vision-based limestone rock-type classification using probabilistic neural network» [2]

В заключении своей работы автор упоминает об этом, он говорит, что основным ограничением данного исследования является то, что использовались цельные куски горных пород, в которые были вкраплены более мелкие образцы. Таким образом, алгоритм будет работать корректно, только тогда, когда тип породы всех мелких включений в образце одинаков.

Автор также отмечает, что данная проблема может быть нивелирована путем сегментации изображения – тогда алгоритм и классификация будут применяться к отдельным включениям в данном изображении.

2. Ручное сопоставление матриц цветовых схем с характеристиками породы

И.И. Хасанов выпустил ряд работ [3]–[8], посвященных полуавтоматической обработке фотографий зерна 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.

3. Применение сверточной нейронной сети (CNN)

Е. Барабошкин, А. Ивченко, Л. Исмаилова и др. в работе «Литологическая интерпретация фотографий зерна на основе компьютерного анализа» [9] приводят результаты исследования, целью которого было создание алгоритма, позволяющего показать тип породы образца зерна. Одна из особенностей работы заключается в сегментации изображения с заданным шагом. Для каждого полученного сегмента вычисляется вероятность, согласно которой выбранная часть образца зерна соотносится с той или иной породой (см. рис. 2). Кроме того, если возможно, определялись дополнительные параметры (нефтенасыщенность, пористость, проницаемость) для данного сегмента зерна.

Для обработки изображений используется библиотека OpenCV [10]. Также в работе применяется алгоритм Dropout [11], который позволяет увеличить точность модели за счет уменьшения влияния переобучения. Производились различные эксперименты с глубиной сети, количеством нейронов на слоях, функциями распространения ошибок, размеров изображений. Сеть показала 84% точности работы в пределах одного класса при размере изображений 128x128 пикселей при равномерном распределении количества образцов.

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 зерна определялись верно, только с точки зрения геологии (тип породы верен) и не верно, с точки зрения машины (класс разрушенный зерн не верно, определён как порода). Количество материалов по классам очень неравномерно: 335 образцов для аргиллита, 43 – для песчаника. Результаты показали до 96 % точности и 98 % полноты.

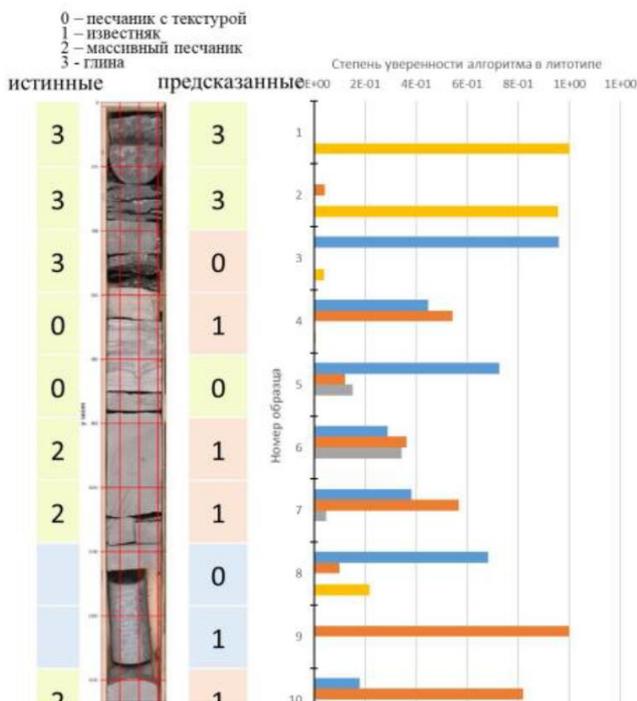


Рис. 2. Пример сегментации образца и определения породы сегмента [9]

брались отметки кровли и подошвы слоя. Далее в рамках одного слоя из изображения последовательно вдоль ствола керна нарезались окна размером 5×5 см с фиксированным шагом смещения для каждого названия породы. Это позволило уменьшить ошибку, возникающую из-за неравномерного распределения пород в обучающей выборке посредством уменьшения шага для менее встречающейся породы в выборке.

В данной работе проводилось сравнение эффективности применения различных подходов для решения поставленной задачи, а именно использовались следующие архитектуры предобученных сверточных нейронных сетей: SqueezeNet1.1, DenseNet121, DenseNet161, Inception-v3, ASNetLarge, PNASNet5Large, ResNet18, ResNet34, ResNet152, ResNeXt101 $64 \times 4d$, ResNeXt101 $32 \times 4d$, SENet154, VGG13, VGG16bn, VGG19bn; применялась технология на ансамбле слабых классификаторов по данным графических идентификаторов. Для реализации данного подхода были использованы две свободно распространяемые библиотеки: Scikit-Image для извлечения графических идентификаторов по окнам и XGBoost для построения модели классификации.

По результатам работы авторами был сделан вывод, что обе модели показали достаточно высокую точность и могут быть применены на практике. Однако модель на основе сверточных нейронных сетей более чувствительна к неявным признакам, которые сложно извлечь, используя известные графические идентификаторы.

Таким образом, проведенный анализ работ по исследованию изображений керна показал:

- одним из основных недостатков предложенных подходов является неравномерность обучающей выборки: при определении типа породы по цветности обучающая выборка состоит из пород преимущественно одного или нескольких литотипов, или и вовсе анализируются не все литотипы; выборка же по остальным литотипам искусственно увеличена;

- в большинстве работ авторы никак не используют изображения образцов керна в ультрафиолетовом свете, хотя ультрафиолетовое свечение позволяет выделить на изображениях нефтенасыщенные участки; выделение таких участков позволит в дальнейшем произвести расчеты нефтенасыщенности и сделать вывод о наличии флюида на том или ином интервале;

- почти все рассмотренные работы сосредоточены на том, чтобы определить тип породы, причем чаще всего на основании цветности. Однако стоит взять во внимание, что съемка без наличия палитры может дать искажения в части цветом, тогда такие изображения будут недостоверными с точки зрения

Однако на некоторых классах (известняк, 124 образца) результаты упали до 42% точности.

На основании полученных данных сделан вывод о возможности использования нейронной сети для определения литотипов образцов керна. В тоже время, полученный продукт имеет определенные недостатки: небольшая обучающая выборка, неравномерность количества литотипов в обучающей выборке, сеть обрабатывает лишь 4 класса пород, низкая точность предсказаний для определенных классов пород, нейросеть работает с большой точностью лишь при соблюдении ряда обязательных условий. Отдельно стоит упомянуть, что данная сеть выполняет описание керна не послойно, а с заданным шагом, что сильно влияет на результаты.

В работе «Автоматическое определение породы по фотографиям керна методами машинного обучения» [1] С. Дьячковым и др. определялась принадлежность слоев керна к одному из 4 типов пород: песчаники, алевролиты, аргиллиты, угли.

Для каждого слоя из исходного описания

определения литотипа по цветности. Кроме того, цветность пород ввиду различных явлений может отличаться, что также не увеличивает достоверность результатов такого метода;

- в некоторых работах при определении типа породы используют либо фиксированный, либо настраиваемый шаг по изображению зерна; при таком методе, если в диапазоне одного шага будет находиться несколько пород, система не сможет выделить обе, а велик шанс, что и вовсе даст неточный результат;

- часть из рассмотренных решений не позволяют производить полную автоматизацию, а лишь упрощают в некоторой степени работу специалиста-геолога.

Описание предложенного алгоритма сверточных нейронных сетей и результаты сегментации образцов зерна на изображениях в дневном свете

На основании выше приведенного анализа, было принято решение использовать сверточные нейронные сети с архитектурой U-Net. Архитектура сети U-Net приведена на рис. 3.

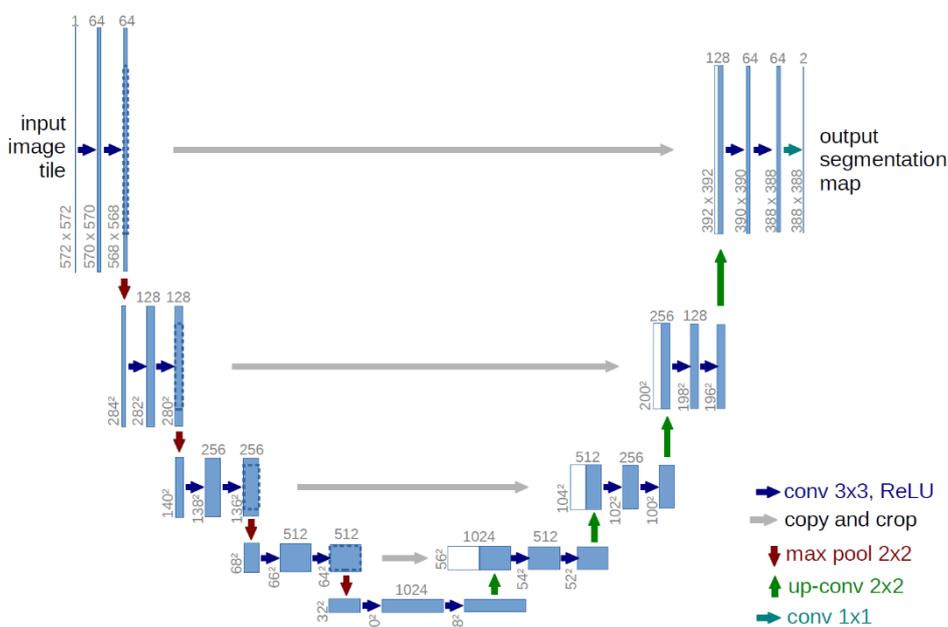


Рис. 3. Архитектура сети U-Net [12]

Она состоит из сужающегося пути (слева) и расширяющегося пути (справа). Сужающийся путь — типичная архитектура сверточной нейронной сети. Он состоит из повторного применения двух сверток 3×3, за которыми следуют ReLU и операция максимального объединения (2×2 степени 2) для понижения разрешения [12].

При этом были сформулированы следующие требования к проведению вычислительного эксперимента: использовать ультрафиолетовые фотографии для более полного анализа; выбирать изображения более высокого качества, чем в аналогичных работах для достижения более точного результата; применять алгоритм Dropout, а также использовать более широкий класс пород на скважинах, с которых отбирается зерно.

Чтобы избежать проблем с неоднородными образцами [2, 9], корректного наложения ультрафиолетовых фотографий для дальнейших расчетов, а также более эффективного обучения и работы дальнейших моделей первым этапом предложенного алгоритма было решено выполнить сегментацию образцов зерна на дневных изображениях. На рис. 4 приведены результаты сегментации образцов зерна на изображениях в дневном свете с применением сверточных нейронных сетей на архитектуре U-Net.

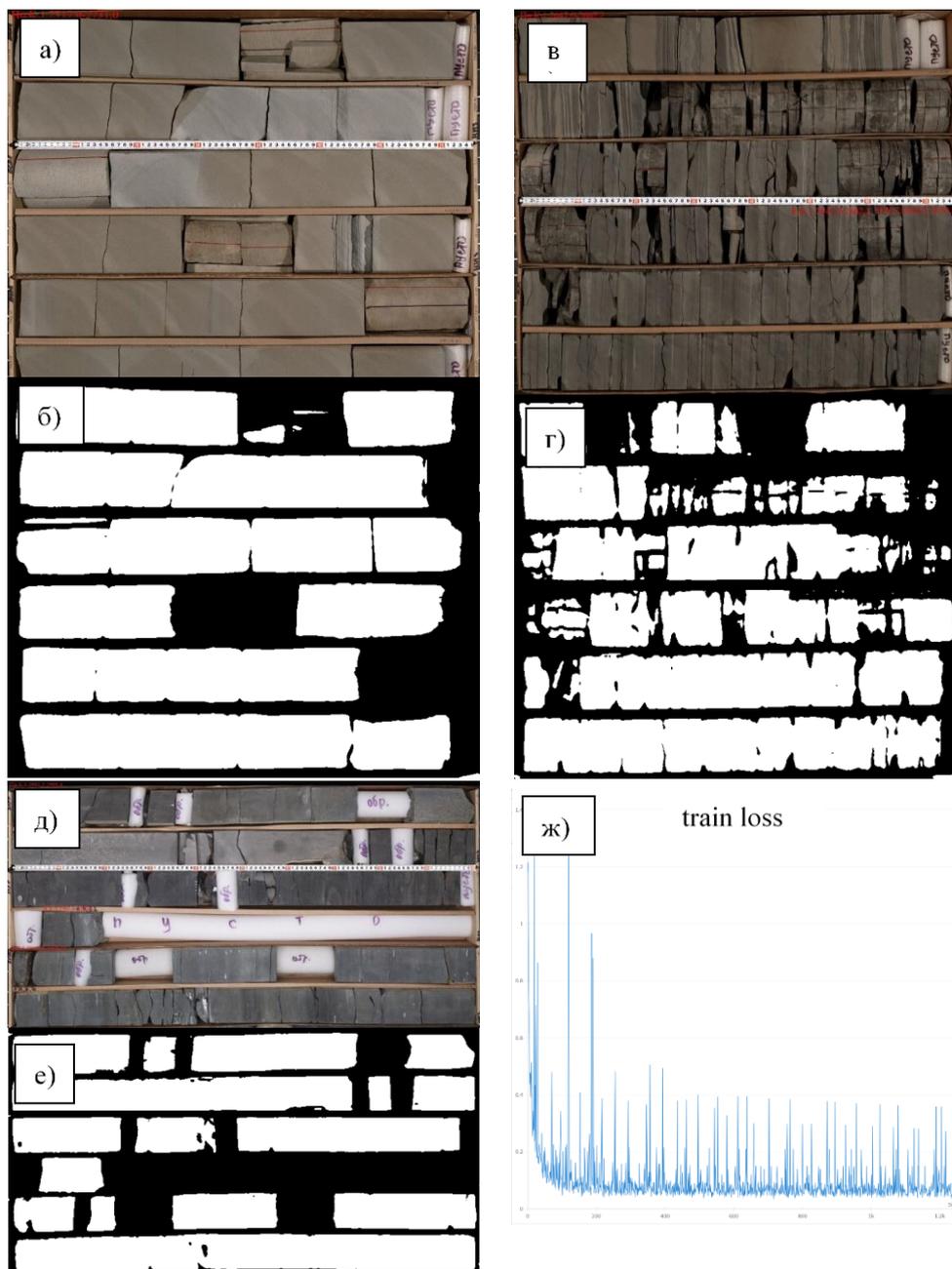


Рис. 4. Примеры сегментации керна на архитектуре U-Net:
 а, в, д – исходные изображения керна; б, г, е – предсказанные маски; ж – функция потерь

На рис. 4 фрагменты а, в, д представляют из себя исходные изображения с образцами керна, а фрагменты б, г, е – маски, предсказанные обученной моделью. Стоит отметить, что на текущий момент модель обучалась на датасете, состоящем в основном из цельной правильной формы образцов керна, что заметно по фрагменту в и г, где модель показала неточные результаты из-за сильной раздробленности образца. Данный фактор говорит о необходимости увеличения датасета в сторону образцов с неправильной формой и большей расчлененностью.

Фрагмент ж представляет собой функцию потерь, которая позволяет наглядно показать эффективность обученной модели.

Заключение

В работе был проведен анализ существующих подходов к исследованию образцов керна. Были отмечены основные преимущества и недостатки каждого из них и на основании сделанных выводов разработан план и требования к проведению дальнейшего исследования образцов керна средствами машинного обучения.

В результате с применением сверточной нейронной сети на архитектуре U-Net была обучена модель для решения задачи сегментации образцов керна на дневных изображениях и приведены результаты работы модели.

В дальнейшем планируется проведение сравнения работы модели на основе архитектуры U-Net с другими архитектурами, расширения обучающей выборки для улучшения качества сегментации на неоднородных образцах породы, а также решение задач сегментации образцов керна на ультрафиолетовых изображениях и ультрафиолетового свечения, определения нефтенасыщенных интервалов.

Список использованных источников

1. С.М. Дьячков, О.А. Ядрышникова, Д.В. Поляков [и др.] Автоматическое определение породы по фотографиям керна современными методами машинного обучения // Вестник Тюменского государственного университета. Физико-математическое моделирование. Нефть, газ, энергетика. – 2021. – Т. 7. – № 4(28). – С. 181-198.
2. Patel A.K., Snehmoy C. Computer vision-based limestone rock-type classification using probabilistic neural network // Geoscience Frontiers. – 2016. – № 7. – С. 53–60.
3. Хасанов И.И. Анализ цветности пород по цифровым изображениям керна // Геология нефти и газа. – 2014. – № 5. – С. 33–39.
4. Постников А.В., Хасанов И.И. Компьютерное моделирование структурно-текстурных особенностей пород // Материалы Балтийской школы-семинара «Петрофизическое моделирование осадочных пород, Петромодел 2015». Санкт-Петербург. 14-18 сентября. 2015.
5. 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.
6. Хасанов И.И. Методика количественной оценки цветности кернового материала и её практическое применение. // Материалы Балтийской школы-семинара «Петрофизическое моделирование осадочных пород, Петромодел 2013» – 2013.
7. 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.
8. 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.
9. Ивченко А.В., Барабошкин Е.Е., Исмаилова Л.С. [и др.] Литологическая интерпретация фотографий керна на основе компьютерного анализа // Международная научная конференция «IEEE Northwest Russia Conference On Mathematical Methods In Engineering And Technology: MMET NW 2018». – Санкт-Петербург: СПбГЭТУ «ЛЭТИ», 2018. – С. 430–433.
10. OpenCV. Modules // Open-Source Computer Vision URL: <https://docs.opencv.org/4.x/> (дата обращения: 27.09.2023).
11. Dropout - метод борьбы с переобучением нейронной сети // Введение в нейронные сети URL: https://proproprogs.ru/neural_network/dropout-metod-borby-s-pereobucheniem-neyronnoy-seti (дата обращения: 24.09.2023).
12. U-Net: нейросеть для сегментации изображений // Виды нейросетей URL: <https://neurohive.io/ru/vidy-nejrosetej/u-net-image-segmentation/> (дата обращения: 20.03.2024).