



Московский государственный университет имени М.В.  
Ломоносова  
Факультет вычислительной математики и кибернетики  
Кафедра исследования операций

Лаврухин Ефим Валерьевич

Применение нейронных сетей для  
сегментации томографических изображений  
геологических пород

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Научный руководитель:  
к.ф.-м.н., доцент  
Д.В. Денисов

Москва, 2018

# Введение

**Актуальность темы исследования.** В настоящее время добыча полезных ископаемых требует большое количество данных о разрабатываемых породах-коллекторах. Эти данные получают в том числе с помощью методов цифровой петрофизики, которые работают с 2-D или 3-D изображениями, сделанными с помощью рентгеновской томографии [1]. Большинство изображений строения пород представлены в грациях серого, которые указывают на интенсивность поглощения рентгеновских лучей. На практике любой метод численного расчета характеристик исходных пород состоит из нескольких отдельных этапов. И первый этап – это сегментация входного изображения, разделение его на несколько различных фаз по плотности вещества. В простейшем случае выполняется бинаризация – разделение на твёрдую породу и поры [2].

**Цель** данной работы – применить методы глубинного обучения, а именно полносвёрточные нейронные сети, для задачи сегментации изображений геологических пород.

**Научная новизна.** В настоящее время существует большое количество методов сегментации. Они существенно отличаются в используемых предположениях о входных изображениях и математическом аппарате. Вот некоторые из них: градиентные [16], морфологические, случайные поля [14], [15], методы Монте-Карло [13]. У этих методов есть ряд достоинств: присутствует математическая формализация, относительная простота постановки задачи, интерпретируемость результатов. Но в то же время все они обладают серьёзным недостатком – в них присутствуют гиперпараметры, которые сильно влияют на качество результата. Это делает затруднительным их применение без оператора, который контролирует процесс и подбирает нужные значения параметров для конкретных

входных данных.

Относительно недавно появились методы сегментации с использованием машинного обучения. Они так же применяются и в сегментации изображений геологических пород [3], [4], [5]. Глубинное обучение - это подкласс моделей машинного обучения, в которых используется сложная многоуровневая композиция слоёв для извлечения нелинейных признаков исходного объекта. Главные преимущества этих моделей - это высокое качество (выше, чем у других методов, решающих аналогичные задачи) и полная автономность обучения (для того, чтобы построить качественную модель на имеющихся данных не требуется участие человека).

Сейчас свёрточные нейронные сети являются, фактически, state-of-the-art в задачах обработки изображений [6] и используются во многих прикладных областях: биологии, медицине, распознавании образов [11], [12]. На текущий момент выпущено достаточно много работ, в которых исследуются методы глубинного обучения для решения связанных с сегментацией пород задач. В частности для построения стохастической реконструкции пород с последующим моделированием физических свойств [7], [8], [9], [10]. Но статей, в которых глубинное обучение применяется для сегментации геологических пород, крайне мало.

В ходе работы для достижения поставленной цели решались следующие **задачи**:

1. Выбор полносвёрточной архитектуры нейронной сети, которая выполняет сегментацию изображений томограмм.
2. Решение проблемы отсутствия размеченных обучающих данных.
3. Построение стабильного алгоритма обучения сети.
4. Учёт внутри модели 3-D структуры входных данных.

5. Сравнение полученных результатов с результатами других моделей для сегментации геологических изображений.
6. Сравнение физических характеристик, полученных с помощью отсегментированных моделью изображений, с характеристиками исходного образца, вычисленными с помощью физических симуляций.

# 1 Постановка задачи

Дано исходное изображение  $S = (s_{ij})_{i=1,j=1}^{H,W}$ ,  $s_{ij} \in [0, 1]$ , где  $H$  - высота изображения,  $W$  - ширина изображения. Требуется для каждого пикселя найти соответствующий ему сегмент изображения, т.е. найти соответствие  $s_{ij} \rightarrow m_{ij}$ ,  $m_{ij} \in C = \{0, \dots, N_c - 1\}$ , где  $N_c$  - число сегментов.

Подразумевается, что для каждого изображения  $S$  существует истинное(возможно, не одно) сегментированное изображение  $M$ . Поэтому можно перейти к следующей постановке задачи: найти алгоритм сегментации  $\mathcal{A}$ , такой, что он преобразует любое изображение  $S$  в маску  $\hat{M}$ :

$$\mathcal{A}(S|\theta) = \hat{M} \quad (1)$$

, где  $\theta$  - настраиваемые параметры нашего алгоритма.

Возникают следующие вопросы:

1. Как выбрать алгоритм  $\mathcal{A}$ ?
2. Как настроить параметры  $\theta$ ?
3. Как оценить ошибку алгоритма?

В данном случае наши параметры  $\theta$  подбираются с помощью обучения без учителя(unsupervised learning), либо непосредственно экспериментатором.

## 2 Особенности задачи сегментации геологических пород

Дан исходный стек изображений  $S = (s_{ijk})_{i=1,j=1,k=1}^{H,W,D}$ ,  $s_{ijk} \in [0, 1]$ , где  $H$  - высота изображения,  $W$  - ширина изображения,  $D$  - количество изображений. Требуется для каждого пикселя найти соответствующую ему метку класса, т.е. найти соответствие  $s_{ijk} \rightarrow m_{ijk}$ ,  $m_{ijk} \in C = \{0, 1\}$ , где класс 0 соответствует порам, класс 1 - твердому веществу.

Особенности данной задачи:

1. Работа с 3-D изображениями.
2. Двухклассовая сегментация.
3. Наличие сегментированных стеков, для которых известно  $\hat{M}$  некоторого "качественного" алгоритма.

Как правило, данная задача на практике решалась методами MRF, snakes и некоторыми другими. Собрана некоторая база изображений  $S, \hat{M}$ , качество сегментации которых оценивалось оператором.

В таком случае естественно перейти к задаче обучения с учителем (supervised learning), чтобы использовать накопившуюся библиотеку сегментированных изображений.

### 3 Постановка задачи supervised сегментации

Дано пространство объектов  $X$  и пространство ответов  $Y$ . Между ними существует соответствие (функция)  $f : X \rightarrow Y$ .

Требуется наилучшим образом приблизить соответствие  $f$  при помощи параметрического семейства функций  $f_\theta$  и множеством примеров отображения  $f : \{(X_1, Y_1), \dots, (X_N, Y_N)\}$ ,  $f(X_i) = Y_i$ ,  $i = \overline{1, N}$ .

В конкретном случае множество примеров из пространства  $X$  задаётся в виде:

$$\hat{X} = \{X_1, \dots, X_N\}, X_k = (x_{ij})_{i=1, j=1}^{H, W}, x_{ij} \in [0, 1] \quad (2)$$

и множество ответов из  $Y$  задаётся в виде:

$$\hat{Y} = \{Y_1, \dots, Y_N\}, Y_k = (y_{ij})_{i=1, j=1}^{H, W}, y_{ij} \in \{0, 1\}. \quad (3)$$

Конкретное значение параметра  $\theta$  параметрического семейства  $f_\theta$  выбирается исходя из функционала качества модели (функции эмпирического риска):

$$Q(\theta, (\hat{X}, \hat{Y})) = \sum_{i=1}^N \mathcal{L}(f_\theta(X_i), Y_i) \longrightarrow \min_{\theta}. \quad (4)$$

Задача состоит в выборе параметрического семейства  $f_\theta$ , функции потерь  $\mathcal{L}(\bar{Y}, Y)$  и решение оптимизационной задачи (4).

## 4 Архитектура нейронной сети

В качестве семейства функций  $f_\theta$  в работе использовалась полносвёрточная нейронная сеть (fully-connected convolutional neutral network). В качестве основы была выбрана архитектура U-net, которая зарекомендовала себя в решении задач биологии (выделение границ клеток).

В архитектуру были внесены незначительные изменения: уменьшено количество свёрточных фильтров, добавлен padding, изменена функция активации на ELU.

Сеть представляет из себя композицию линейных (свертки (7), конкатенация (8), и transposed convolutions, действие которых эквивалентно обратному действию свёрточных слоев, т.е по выходу свёрточного слоя  $y$  и фильтрам  $w$  получается вход свёрточного слоя  $x$ ) и нелинейных преобразований (активации (5), (6), pooling (9)) которые применяются последовательно. Вероятности на выходе обеспечиваются с помощью softmax-преобразования выхода сети:

$$y = \begin{cases} x, & \text{при } x \geq 0, \\ e^x - 1, & \text{при } x < 0. \end{cases} \quad (5)$$

$$y_{c,i,j} = \frac{e^{x_{c,i,j}}}{\sum_{l=1}^C e^{x_{l,i,j}}} \quad (6)$$

$$x \in \mathbb{R}^{C \times W \times H}, \quad y \in \mathbb{R}^{C \times W \times H}$$



$$\begin{aligned}
y_{c,i,j} &= \sum_{l=1}^C \sum_{m=1}^{\min(W-i,K)} \sum_{n=1}^{\min(H-j,K)} x_{l,i+m,j+n} w_{l,m,n}^c \\
, \quad c &= \overline{1,T}, \quad i = \overline{1,W}, \quad j = \overline{1,H} \\
, \quad \text{где } x &\in \mathbb{R}^{C \times W \times H} - \text{вход свёртки} \\
, \quad y &\in \mathbb{R}^{T \times W \times H} - \text{выход свёртки} \\
, \quad w^c &\in \mathbb{R}^{C \times K \times K} - \text{фильтры свёрток} \\
, \quad c &= \overline{1,C} - \text{количество свёрток.}
\end{aligned} \tag{7}$$

$$z_{c,i,j} = \begin{cases} x_{c,i,j}, & \text{если } c \leq C, \\ y_{c-C,i,j}, & \text{если } c > C \end{cases} \tag{8}$$

, где  $x \in \mathbb{R}^{C \times W \times H}$ ,  $y \in \mathbb{R}^{T \times W \times H}$ ,  $z \in \mathbb{R}^{(T+C) \times W \times H}$

$$\begin{aligned}
y_{c,i,j} &= \max_{\substack{m=1, \min(W-i,K) \\ n=1, \min(H-j,K)}} x_{c,iK+m,jK+n} \\
, \quad \text{где } x &\in \mathbb{R}^{C \times W \times H} - \text{вход pooling'a} \\
, \quad y &\in \mathbb{R}^{C \times \left\lceil \frac{W+K-1}{K} \right\rceil \times \left\lceil \frac{H+K-1}{K} \right\rceil} - \text{выход pooling'a} \\
, \quad K &- \text{размер ядра pooling'a.}
\end{aligned} \tag{9}$$

Модель реализует следующее отображение:

$$\begin{aligned}
f_{\theta}(x) &= y, \\
x &\in \mathbb{R}^{1 \times H \times W}, \quad y \in \mathbb{R}^{2 \times H \times W}
\end{aligned} \tag{10}$$

, где первая первая размерность - число каналов изображения (на входе один, потому что изображение grayscale; на выходе два: в первом канале вероятность того, что текущий пиксель - это пора, во втором - что это твёрдое вещество), последние две размерности - это размер изображений. В качестве оптимизируемых параметров модели  $\theta$  выступает

совокупность весов convolutional и transposed convolutional слоёв.

## 5 Оптимизация модели

1

## 6 Заключение

### Список литературы

- [1] S. Karimpoulia , P. Tahmasebib, , H. L. Ramandic , P. Mostaghimid , M. Saadatfar, “Stochastic modeling of coal fracture network by direct use of microcomputed tomography images”, International Journal of Coal Geology 179, 153-163, 2017.
- [2] Jeff T. Gostick, “Versatile and efficient pore network extraction method using marker-based watershed segmentation”, Physical Review E 96, 2017.
- [3] S. Chauhan , W. Rühaak, , H. Anbergen , A. Kabdenov , M. Freise , T. Wille , I. Sass, “Phase segmentation of X-ray computer tomography rock images using machine learning techniques: an accuracy and performance study”, Solid Earth, 7, 1125–1139, 2016.
- [4] F. Khan , F. Enzmann, , M. Kersten, “Multi-phase classification by a least-squares support vector machine approach in tomography images of geological samples”, Solid Earth, 7, 481–492, 2016.
- [5] S. Chauhan , W. Rühaak, , F. Khan , F. Enzmann , P. Mielke , I. Sass, “Processing of rock core microtomography images: Using seven different machine learning algorithms”, Computers & Geosciences, 86, 120-128, 2016.
- [6] O. Ronneberger , P. Fischer , T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation”, arXiv:1505.04597v1, 2015.
- [7] L. Mosser , O. Dubrule, , Martin J. Blunt, “Reconstruction of three-

dimensional porous media using generative adversarial neural networks”, arXiv:1704.03225v1, 2017

- [8] Brian L. DeCost , T. Francis , Elizabeth A. Holm, “Exploring the microstructure manifold: image texture representations applied to ultrahigh carbon steel microstructures”, arXiv:1702.01117v2, 2017
- [9] Ruijin Cang , Yaopengxiao Xu , Shaohua Chen , Yongming Liu , Yang Jiao ,M. Yi Ren, ‘Microstructure Representation and Reconstruction of Heterogeneous Materials via Deep Belief Network for Computational Material Design”, arXiv:1612.07401v3, 2017
- [10] N. Lubbers , T. Lookman , K. Barros, “Inferring low-dimensional microstructure representations using convolutional neural networks”, arXiv:1611.02764v1, 2016
- [11] A. S. Razavian , H. A. Josephine , S. S. Carlsson, “CNN Features off-the-shelf: an Astounding Baseline for Recognition”, arXiv:1403.6382v3, 2014.
- [12] Kwang Moo Yi , Eduard Trulls , Vincent Lepetit , Pascal Fua, “LIFT: Learned Invariant Feature Transform”, arXiv:1603.09114v2, 2016.
- [13] С.А. Эль-Хатиб, “Сегментация изображений с помощью смешанного и экспоненциального алгоритмов роя частиц”, Информатика и кибернетика, 1, 2015.
- [14] H. Deng , D.A. Clausi, “Unsupervised image segmentation using a simple MRF model with a new implementation scheme”, Pattern Recognition, 37, 2323-2335, 2004.

- [15] H. Deng , D.A. Clausi, “Image segmentation using Markov Random Field Model in Fully Parallel Cellular Network Architecture”, *Real-time Imaging*, 6, 195-211, 2000.
- [16] H. Deng , D.A. Clausi, “Improved Workflow for Unsupervised Multiphase Image Segmentation”, *arXiv:1710.0967*, 2017.