



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

Лаврухин Ефим Валерьевич

**Применение нейронных сетей для решения задачи  
оптимальной сегментации томографических  
изображений геологических пород**

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

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

Москва, 2018

# Содержание

<b>1</b>	<b>Введение</b>	<b>2</b>
<b>2</b>	<b>Сегментации изображений геологических пород</b>	<b>4</b>
2.1	Постановка задачи сегментации . . . . .	4
2.2	Задача сегментации геологических пород . . . . .	5
2.3	Постановка задачи сегментации с учителем . . . . .	6
<b>3</b>	<b>Выбор модели</b>	<b>7</b>
<b>4</b>	<b>Оптимизация модели</b>	<b>10</b>
4.1	Оптимизируемый функционал . . . . .	10
4.2	Метод оптимизации . . . . .	11
<b>5</b>	<b>Ход работы</b>	<b>13</b>
5.1	Прикладная задача . . . . .	13
5.2	Данные для экспериментов . . . . .	13
5.3	Разделение данных на обучающие и тестовые . . . . .	13
<b>6</b>	<b>Результаты</b>	<b>15</b>
6.1	Обучение моделей . . . . .	15
6.2	Выбор порога вероятности . . . . .	15
6.3	Оценка ошибок в разных плоскостях . . . . .	19
6.4	Оценка качества моделей . . . . .	19
<b>7</b>	<b>Заключение</b>	<b>24</b>
	<b>Литература</b>	<b>25</b>

# 1 Введение

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 градациях 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].

Существует большое количество методов сегментации. Они существенно отличаются в принимаемых предположениях о входных изображениях и используемом математическом аппарате: градиентные [13], морфологические [16], случайные поля [14], методы Монте-Карло [15]. 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. Классические методы машинного обучения уже применялись в сегментации изображений геологических пород [3], [4], [5].

Сейчас свёрточные нейронные сети являются, фактически, state-of-the-art в задачах обработки изображений и используются во многих прикладных областях: биологии [6], медицине [17], распознавании образов [11]. 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.

Цель данной работы – построить модель для сегментации томографических изображений геологических пород, которая бы имела высокое качество работы и не была бы чувствительной к настройке оператором. В данной работе модель будет строиться с помощью методов глубинного обучения, которые, как говорилось ранее, зарекомендовали себя в похожих задачах. В качестве модели будет использоваться полносвёрточная нейронная сеть.

В ходе работы решались следующие задачи:

1. Постановка задачи оптимальной сегментации как задачи нелинейной оптимизации(гл. 2).
2. Выбор полносвёрточной архитектуры нейронной сети для выполнения сегментацию изображений томограмм(гл. 3).
3. Построение стабильного алгоритма обучения сети(гл. 4).
4. Решение проблемы отсутствия размеченных обучающих данных(гл. 5).
5. Проверка модели на реальных данных(гл. 6).
6. Анализ полученных результатов(гл. 7).

## 2 Сегментации изображений геологических пород

В данной главе будет формально поставлена практическая задача сегментации томографических изображений геологических пород: самая общая постановка (п. 2.1) и постановки, которые используются в данной работе (п. 2.2, 2.3).

### 2.1 Постановка задачи сегментации

Рассмотрим простейшую постановку задачи сегментации изображения.

#### Задача сегментации изображения

**Дано:** изображение  $S = (s_{ij})_{i=1,j=1}^{H,W}$ ,  $s_{ij} \in [0, 1]$ , где  $H$  – высота изображения,  $W$  – ширина изображения.

**Найти:** для каждого пикселя соответствующий ему сегмент изображения, т.е. соответствие  $s_{ij} \rightarrow m_{ij}$ ,  $m_{ij} \in C = \{0, \dots, N_c - 1\}$ , где  $N_c$  – число различных сегментов.

Подразумевается, что для каждого изображения  $S$  существует истинное (возможно, не одно) сегментированное изображение  $M = (m_{ij})_{i=1,j=1}^{H,W}$ . 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.

Для задачи возникают следующие вопросы:

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

Существуют различные способы решения этих вопросов. Выбор алгоритма  $\mathcal{A}$  и способа оценки качества его работы обуславливается задачей и необходимыми результатами. Подбор параметров  $\theta$  выполняется непосредственно оператором, с помощью обучения без учителя (unsupervised learning) или с помощью обучения с учителем (supervised learning).

## 2.2 Задача сегментации геологических пород

Для задачи сегментации томографических изображений геологических пород существуют некоторые особенности.

### Задача сегментации томографических изображений

**Дано:** 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.

**Найти:** для каждого пикселя соответствующую ему метку класса, т.е. соответствие  $s_{ijk} \rightarrow m_{ijk}$ ,  $m_{ijk} \in C = \{0, 1\}$ , где класс 0 соответствует полостям в веществе, а класс 1 – твердому веществу.

Особенности данной задачи, по сравнению с аналогичной в параграфе 2.1:

1. Работа с 3-D изображениями.
2. Двухклассовая сегментация.
3. Наличие обработанных образцов, для которых известно  $\hat{M}$  некоторого алгоритма, признанного эталонным.

Данная задача является задачей обучения без учителя (unsupervised learning), поэтому методы, которые работают в рамках такой постановки задачи, не используют какую-либо апостериорную информацию об обрабатываемом образце. Как правило, данная задача на практике решается методами MRF [14], САС [12] и некоторыми другими. Эти методы обладают набором внешних параметров (аналогично с алгоритмом (1)) и чувствительны к их изменению.

На текущий момент существуют коллекции пар образец-сегментация  $(S, \hat{M})$ , полученных с помощью методов без обучения с параметрами, подобранными оператором исходя из своей экспертизы. Качество выборок, полученных с помощью эталонных методов (1), можно считать достаточно хорошим для того, чтобы естественно перейти к задаче обучения с учителем (supervised learning), использовав накопившуюся библиотеку сегментированных 3-D изображений.

## 2.3 Постановка задачи сегментации с учителем

Поставим задачу сегментации с учителем.

### Задача обучения с учителем

**Дано:** пространство объектов  $X$  и пространство ответов  $Y$ . Между ними существует соответствие (функция)  $f : X \rightarrow Y$ , множество примеров отображения  $f : \{(X_1, Y_1), \dots, (X_N, Y_N)\}$ ,  $f(X_i) = Y_i$ ,  $i = \overline{1, N}$ , дано параметрическое семейство функций  $f_\theta : X \rightarrow Y$ .

**Найти:** наилучшим образом приблизить соответствие  $f$  на всём пространстве  $X$  с помощью параметрического семейства функций  $f_\theta$ , используя данное множество примеров отображения.

В конкретном случае пространство  $X$  является пространством изображений, это накладывает ограничения на виде объектов. Множество примеров из пространства  $X$  задаётся в виде:

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

Пространство ответов из  $Y$  так же является пространством изображений, но с дискретными элементами. Сейчас рассматривается бинарная сегментация, поэтому множество примеров задаётся в виде:

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

Задача, по сути, сводится к выбору параметра  $\theta$  в рамках параметрического семейства  $f_\theta$ . Конкретное значение параметра обычно выбирается исходя из функционала качества модели (функции эмпирического риска):

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

Обычно выбирается подходящая по свойствам функция потерь  $\mathcal{L}(\bar{Y}, Y)$ , и оптимизационная задача (4) решается с помощью какого-либо метода оптимизации: метода первого порядка (стохастического градиента), метода второго порядка (LBFGS), максимизации правдоподобия (ЕМ), эвристического (отжиг) или другого.

### 3 Выбор модели

В качестве семейства функций  $f_\theta$  для аппроксимации зависимости из задачи параграфа 2.3 в данной работе использовалась полносвёрточная нейронная сеть (fully-connected convolutional neutral network). В качестве основы была выбрана архитектура U-net [6], которая зарекомендовала себя в решении задач биологии(выделение границ клеток).

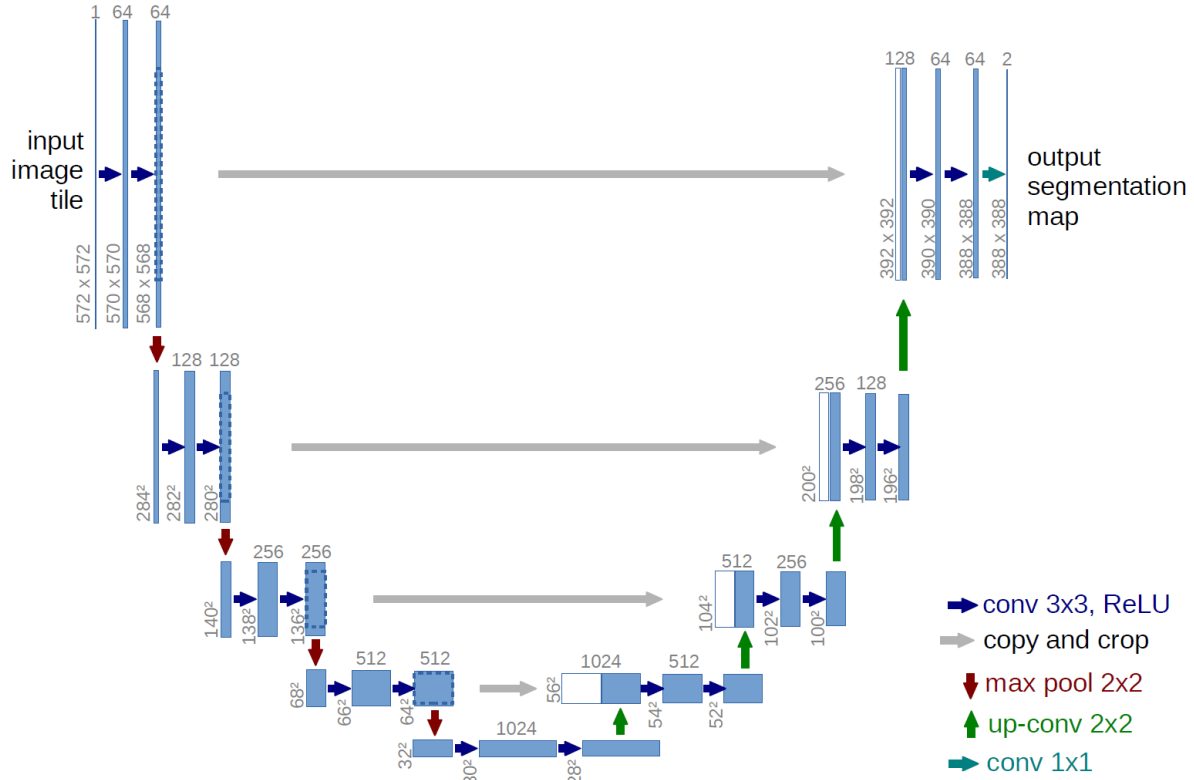


Рис. 1: Иллюстрация архитектуры U-net из оригинальной статьи [6]

Сеть была названа таким образом из-за U-образной формы. Часть от входа до самого узкого слоя посередине называется кодировщик(encoder), она служит для сжатия входного изображения и извлечения из него самых существенных для решаемой задачи признаков. Часть от середины и до выхода называется декодировщик(decoder), она служит для восстановления основной информации исходного изображения(задачу сегментации можно интерпретировать как удаление шума из исходного изображения).

В архитектуру [6] были внесены незначительные изменения:

- Уменьшено количество свёрточных фильтров (с 64 до 32).



- Добавлено заполнение по краям свёрток(padding).
- Изменена функция активации(с ReLU на ELU).

Сеть представляет из себя композицию следующий слоёв:

- Линейных свертков (5):

$$\begin{aligned}
y_{c,i,j}(x|w) &= \sum_{l=1}^C \sum_{m=1}^{\min(W-i,K)} \sum_{n=1}^{\min(H-j,K)} x_{l,i+m,j+n} w_{l,m,n}^c \\
, \quad c &= \overline{1,T}, \quad i = \overline{1,W}, \quad j = \overline{1,H} \\
, \text{ где } x &\in \mathbb{R}^{C \times W \times H} - \text{ вход свёртки} \\
, y &\in \mathbb{R}^{T \times W \times H} - \text{ выход свёртки} \\
, w^c &\in \mathbb{R}^{C \times K \times K} - \text{ фильтры свёрток} \\
, c &= \overline{1,C} - \text{ количество свёрток.}
\end{aligned} \tag{5}$$

- Объединений выходов слоёв (concatenation, 6):

$$\begin{aligned}
z_{c,i,j}(x,y) &= \begin{cases} x_{c,i,j}, & \text{если } c \leq C, \\ y_{c-C,i,j}, & \text{если } c > C \end{cases} \\
, \text{ где } x &\in \mathbb{R}^{C \times W \times H}, \quad y \in \mathbb{R}^{T \times W \times H}, \quad z \in \mathbb{R}^{(T+C) \times W \times H}
\end{aligned} \tag{6}$$

- "Обратных" свёрток(transposed convolution, действие которых эквивалентно обратному действию свёрточных слоев, т.е по выходу свёрточного слоя  $y$  и фильтрам  $w$  получается вход свёрточного слоя  $x$ ).
- Нелинейных преобразований активации (7), (8), pooling (9).

$$y(x) = \begin{cases} x, & \text{при } x \geq 0, \\ e^x - 1, & \text{при } x < 0. \end{cases} \tag{7}$$

$$\begin{aligned}
y_{c,i,j}(x) &= \frac{e^{x_{c,i,j}}}{\sum_{l=1}^C e^{x_{l,i,j}}} \\
x &\in \mathbb{R}^{C \times W \times H}, \quad y \in \mathbb{R}^{C \times W \times H}
\end{aligned} \tag{8}$$

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

- Слой выпадения сигнала(dropout (10)).

$$y(x) = D \odot x, \text{ где } x \in \mathbb{R}^{C \times W \times H}, D = (d_{kijk})_{k=1, i=1, j=1}^{C, W, H}, d_{ijk} \sim B(p) \tag{10}$$

Данные преобразования применяются последовательно, порядок композиции показан на рис.1. Количество pooling(9), concatenation(6) и transposed convolution слоёв подобрано так, чтобы выходной тензор имел такие же пространственные размерности  $H \times W$ , что и входной, и 2 канала(вместо одного на входе). Для улучшения обобщающей способности сети и предотвращению переобучения на стыке кодировщика и декодировщика расположен слой (10). Вероятности на выходе обеспечиваются с помощью softmax-преобразования выхода сети (8).

В итоге модель реализует следующее отображение:

$$\begin{aligned}
f_{\theta}(x) &= y, \\
x &\in \mathbb{R}^{1 \times H \times W}, y \in \mathbb{R}^{2 \times H \times W}
\end{aligned} \tag{11}$$

, 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 в задаче (4) параметров модели  $\theta$  выступает совокупность весов convolutional и transposed convolutional слоёв.

## 4 Оптимизация модели

В данной главе даётся законченная формулировка задачи (4), вводится конкретная функция потерь. Указывается метод оптимизации для решения поставленной задачи с конкретной моделью из предыдущей главы.

### 4.1 Оптимизируемый функционал

Для финальной постановки задачи осталось выбрать функцию потерь  $\mathcal{L}(\bar{Y}, Y)$ . Поскольку выход модели (11) является бинарными вероятностями, подходящей функцией является кросс-энтропия:

$$CE(\bar{Y}, Y) = -\frac{1}{WH} \sum_{i=1, j=1}^{W, H} \left( Y_{ij} \log \bar{Y}_{0ij} + (1 - Y_{ij}) \log(1 - \bar{Y}_{1ij}) \right). \quad (12)$$

Выбор обуславливается тем, что функция 12 гладкая, выпуклая, минимум достигается при выборе с помощью прогноза  $\bar{Y}$  верного класса, и она имеет понятную и адекватную вероятностную интерпретацию.

Так же в качестве метрики качества в задачах сегментации используют коэффициент Жаккара:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (13)$$

, где в качестве множеств  $A$  и  $B$  выступают множество пикселей, отнесённые к первому(второму) классу моделью, и множество пикселей, образующих правильный ответ для первого(второго) класса.

Коэффициент Жакарра (13) нельзя использовать для обучения модели напрямую, потому что для его вычисления требуется определённый(метка конкретного класса) прогноз модели, а модель (11) возвращает вероятности меток. При переходе от вероятности к меткам по порогу(например, при  $p < \tau = 0.5$  предсказывается первый класс, иначе – второй) функция перестает быть гладкой:

$$y(x) = \begin{cases} 0, & \text{при } x \leq \tau, \\ 1, & \text{при } x > \tau. \end{cases} \quad (14)$$

Поэтому в качестве функции потерь можно использовать гладкую аппроксимацию (13). Можно придумать различные виды аппроксимации, в данной работе при-

менялась следующая(такой подход использован в [19]):

$$sIOU(\bar{Y}, Y) = \frac{1}{WH} \sum_{i=1, j=1}^{W, H} \frac{Y_{ij} \bar{Y}_{ij} + \varepsilon}{Y_{ij} + \bar{Y}_{ij} - Y_{ij} \bar{Y}_{ij} + \varepsilon}. \quad (15)$$

Можно убедиться, что при "стремлении"  $\bar{Y}$  к  $Y$  значение  $sIOU(\bar{Y}, Y)$  стремится к  $J(A(\bar{Y}, \tau), Y)$  при любом выборе порога  $\tau$ . В выражении (15) фигурирует константа  $\varepsilon$  (на практике, например,  $\varepsilon = 10^{-5}$ ), которая используется для вычислительной стабильности.

Итоговая функция потерь, которая использовалась для оптимизации модели, имеет вид:

$$\mathcal{L}(\bar{Y}, Y) = CE(\bar{Y}, Y) - \gamma \log(sIOU(\bar{Y}, Y)) \quad (16)$$

, где  $\gamma$  - коэффициент соотношения функций потерь. Величина (15) находится под логарифмом для коррекции соотношения к величине (12). Аппроксимация для индекса Жакарра взята с отрицательным знаком потому, что при увеличении качества модели значение индекса увеличивается.

## 4.2 Метод оптимизации

Последним шагом для завершения модели является выбор метода решения задачи (11) для функции потерь (16).

Для задач оптимизации параметров нейронных сетей характерны следующие особенности:

- Большое число оптимизируемых параметров.
- Непрерывное пространство оптимизируемых параметров.
- Возможность аналитического вычисления градиентов с помощью алгоритма обратного распространения ошибок(back propogation).
- Сложная структура оптимизируемого функционала: отсутствие выпуклости, множество седловых точек и локальных минимумов.
- Большой размер обучающей выборки.

В связи с непрерывной структурой пространства параметров и возможностью вычислять аналитические градиенты, эвристические(отжиг, генетические алгоритмы) методы для оптимизации параметров нейронных сетей применяются редко. Из-за огромного числа параметров квадратичные методы оптимизации не применимы. Квазиньютоновские методы тоже, как правило, оказываются неоправданно медленными. Для обучения используются тренировочные выборки большого размера, поэтому методы оптимизации должны обрабатывать выборку порционно(по батчам).

По этим причинам задачи оптимизации в нейронных сетях решаются в основном с помощью вариаций стохастического градиента. В основном используются варианты с моментумом(для выхода из седловых точек и плохих локальных минимумов) и адаптивным темпом обучения. В данной работе в качестве метода оптимизации использовался ADAM [18].

## 5 Ход работы

### 5.1 Прикладная задача

Задача, которая стояла перед началом работы, – проверить возможность применения глубинного обучения для сегментации изображений геологических пород. В настоящее время для решения данной задачи используются методы без обучения, поэтому для настройки параметров методов необходим оператор, который будет валидировать полученные результаты. Методы глубинного обучения выглядят привлекательными для данной задачи по следующим причинам:

- Автоматическая оптимизация параметров, а значит полной независимости от оператора.
- Оптимизация на этапе обучения, отсутствие оптимизации на этапе прогноза.
- Высокая обобщающая способность, возможность построения универсальной модели для широкого класса пород.

### 5.2 Данные для экспериментов

Из-за наличия обучения возникает проблема формирования обучающей выборки: сейчас не существует коллекции правильно размеченных томограмм геологических пород. Альтернативой является использование в качестве обучающей выборки изображений, полученных с помощью других методов под контролем оператора. Именно такой путь был выбран в данной работе.

В качестве тренировочной выборки были выбраны образцы следующих пород: карбонаты(4 образца, названия: carb96558, carb71, carbRNF, SPE-carb) песчаники(4 образца, названия: SoilAh-1, SoilB-1, SoilB-2, TiTree-1, TiTree-2) и почвы(3 образца, названия: Urna-22, Urna-30, Urna-34). В качестве истинной сегментации использовались изображения, полученные с помощью метода CAC(convergene active contours) [12].

### 5.3 Разделение данных на обучающие и тестовые

В ходе работы было построено 7 моделей, каждая модель обучалась на 700 парах изображение-сегментация, изображения размера  $700 \times 700$ . Модели можно разделить

на три группы:

1. Одиночные модели:

- Модель для карбонатов обучалась на образце carb96558.
- Модель для почв обучалась на образце SoilB-2.
- Модель для песчаников обучалась на образце Urna-22.

2. Парные модели:

- Парная модель для карбонатов и почв обучалась на образцах carb96558 и SoilB-2, обучающие пары распределялись в пропорции 1 : 1.
- Парная модель для карбонатов и песчаников обучалась на образцах carb96558 и Urna-22, обучающие пары распределялись в пропорции 1 : 1.
- Парная модель для песчаников и почв обучалась на образцах Urna-22 и SoilB-2, обучающие пары распределялись в пропорции 1 : 1

3. Общая модель для песчаников, почв и карбонатов обучалась на образцах Urna-22, SoilB-2 и carb96558, обучающие пары распределялись в пропорции 1 : 1 : 1.

В качестве образцов для обучения намерено выбирались наименее "особенные" образцы, чтобы проверить обобщающую способность модели, возможности справляться с объектами с отличиями от обучающей выборки.

## 6 Результаты

### 6.1 Обучение моделей

Процесс обучения состоял из 10 эпох. В каждой эпохе модель обучалась на случайной подвыборке тренировочной выборки, размер случайной подвыборки  $\frac{1}{5}$  от всей выборки. Таким образом каждый обучающий пример подавался в модель в среднем 2 раза за время обучения. В качестве входных данных модели подавались 2-D фрагменты изображений образцов размером  $64 \times 64$ , которые были получены разделением исходных изображений на части с минимальным наложением. Все изображения образца расположены в одной плоскости, 3-D структура образца формируется путём "наложения" изображений друг на друга.

В качестве целевого функционала использовался (16) с  $\gamma = 1, \varepsilon = 10^{-5}$ . Модель обучалась с помощью метода ADAM [18] с параметрами  $\alpha = 10^{-3}, \beta_1 = 0.9, \beta_2 = 0.999$ . Так же использовалось мультипликативное понижение темпа обучение, т.е. на каждой эпохе  $\alpha$  домножалось на темп снижения  $\lambda = \sqrt[10]{10^{-2}}$  (такая константа выбрана для того, чтобы на последней эпохе темп обучения снизился до  $\alpha = 10^{-5}$ ).

В процессе обучения отслеживалась величина оптимизируемого функционала (16), а так же следующие функции качества: logistic loss (12), sIOU (15) и точность предсказаний по фиксированному порогу  $\tau = 0.5$ .

На рис. 2, 3 приведены примеры кривой обучения для нескольких моделей. Видно, что метрики ведут себя согласованно (т.е. сохраняется тренд повышения качества по всем метрикам). К последним эпохам метрики выходят на плато, что говорит о стабилизации параметров модели.

### 6.2 Выбор порога вероятности

В рамках прикладной задачи по обработке образцов томограмм геологических пород от модели требуется получение бинарной маски предсказанных классов для пикселей исходного изображения. Модель (11) предоставляет распределение вероятностей классов. Для того, чтобы перейти от вероятностей к бинарным меткам, требуется выбрать порог перехода  $\tau$  (для отслеживания прогресса в обучении модели, как указано в предыдущем подразделе, выбиралось  $\tau = 0.5$ ). Поскольку конечная



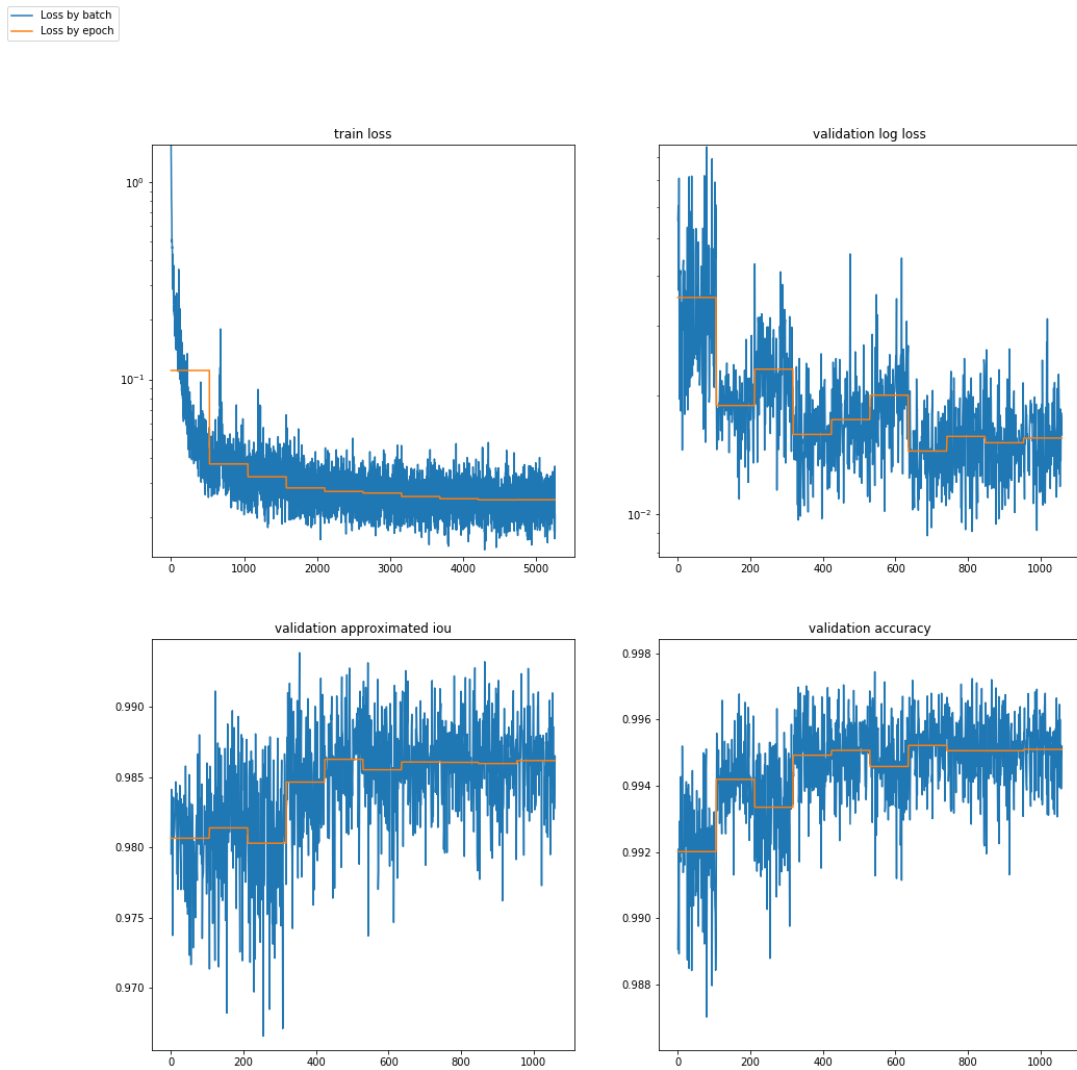


Рис. 2: График кривой обучения для одиночной модели из группы 1, обученной на образце SoilB-2.

цель данной работы - это получение модели с высокой обобщающей способностью, применимой к различным видам пород, подбор параметра порога  $\tau$  выглядит затруднительной задачей (в самом деле, он может быть разным для разных образцов пород, а значит его нельзя заранее получить для образцов, которые обрабатываются впервые, и для которого нет аналогичных обработанных образцов).

В ходе экспериментов было установлено, что модель (11), которая тренируется для минимизации функции качества (16), имеет необычную (нехарактерную для подобных моделей) плотность распределения вероятностей пикселей  $f(p)$  - вся вероятностная масса расположена вблизи 0 и 1. Это значит, что качество бинарного

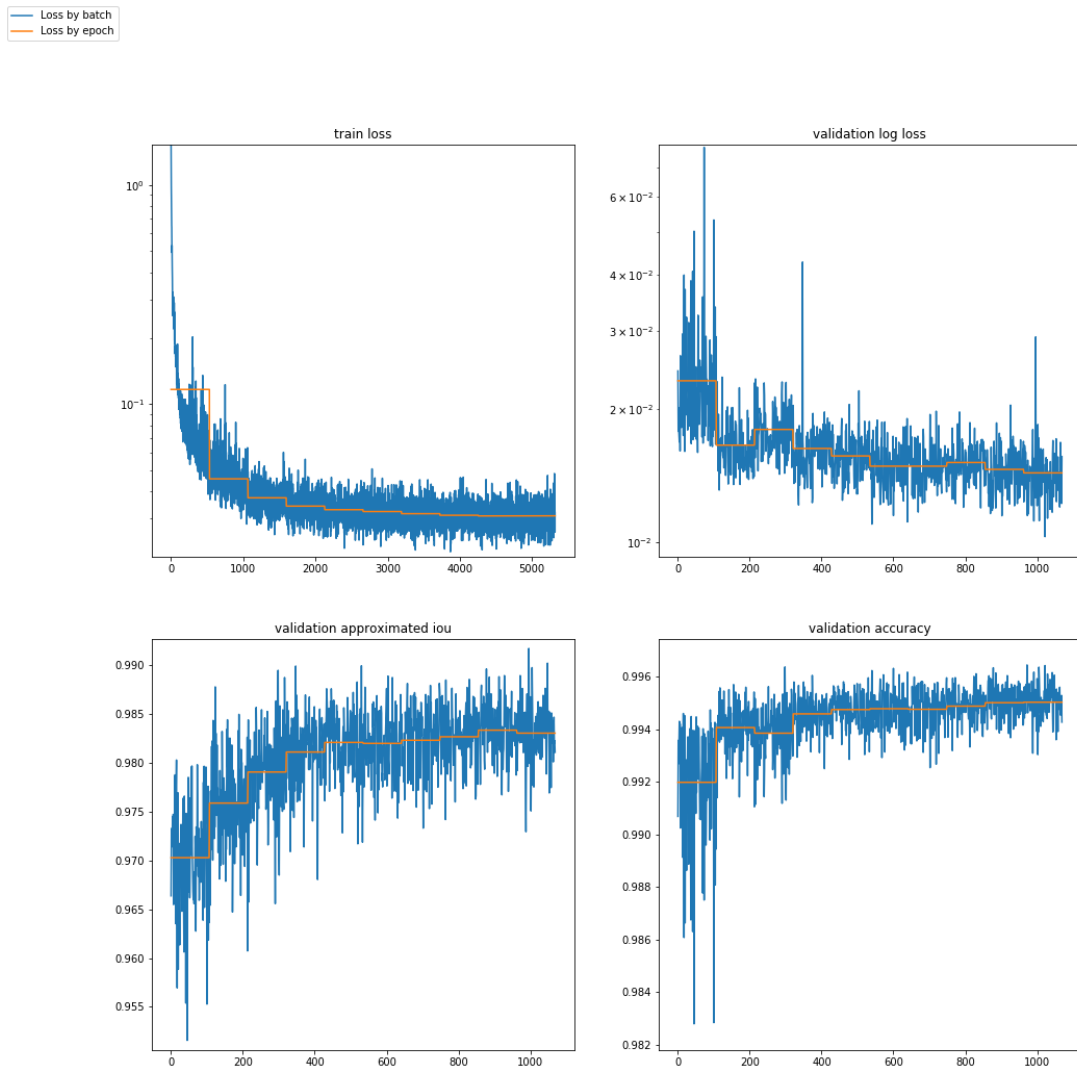


Рис. 3: График кривой обучения для общей модели группы 3.

прогноза модели устойчиво к выбору порога в широком интервале(в эксперимента любой порог  $\tau \in [0.1, 0.9]$  показывал примерно аналогичный результат, иллюстрация на рис. 4). Из вышесказанного следует, что данное свойство модели имеет большое практическое значение.

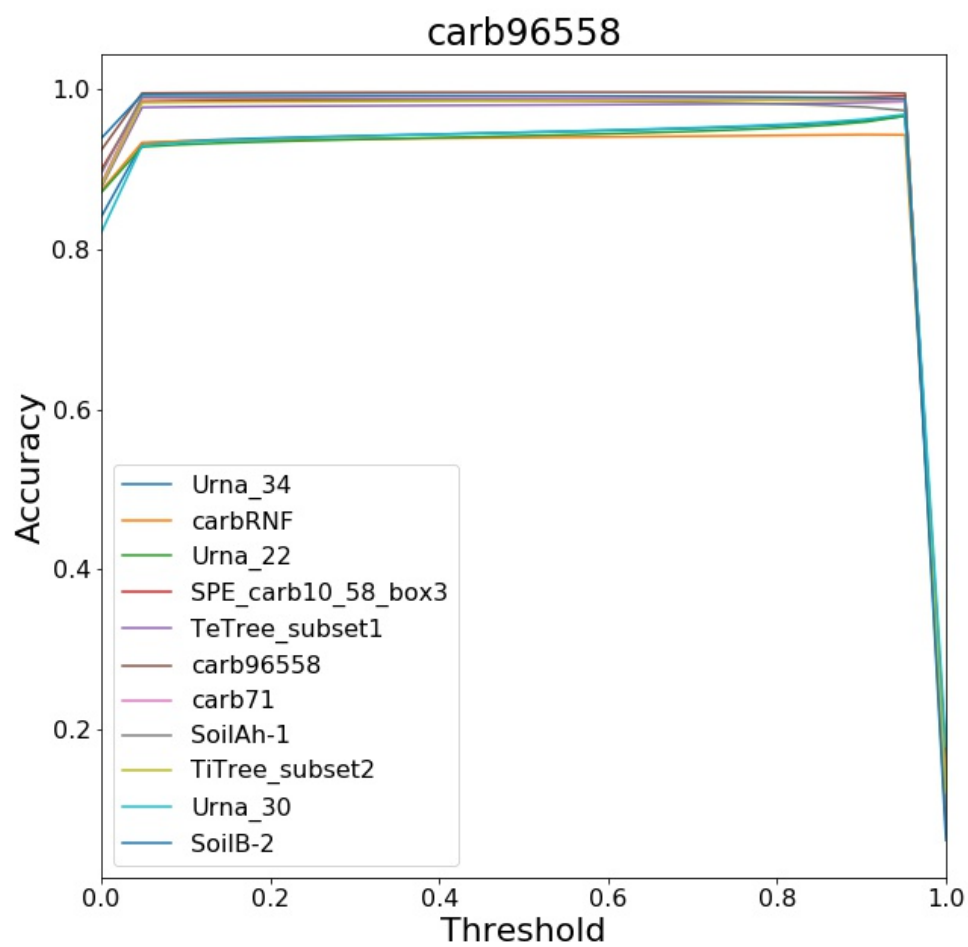


Рис. 4: Зависимость точности предсказания от величины порога вероятности  $\tau$  для модели из группы 1, обученной на образце carb96558.

### 6.3 Оценка ошибок в разных плоскостях

Поскольку обучающие примеры были взяты из одной плоскости, а образцы являются трёхмерными, требовалось проверить качество предсказания модели в другой плоскости, ортогональной плоскости, из которой брались обучающие примеры.

Эксперименты показали, что нет существенного различия в качестве сегментации для изображений, взятых из различных плоскостей, имеет место равенство функционалов качества, а так же визуальная эквивалентность результатов.

Примеры полученных результатов на рис. 5, 6. Зелёным цветом показанны правильно предсказанные моделью поры, красным – неправильно предсказанные моделью поры, синим – неправильно предсказанное моделью твёрдое вещество. Видно, что предсказание модели и правильный ответ для изображений почти не различается.

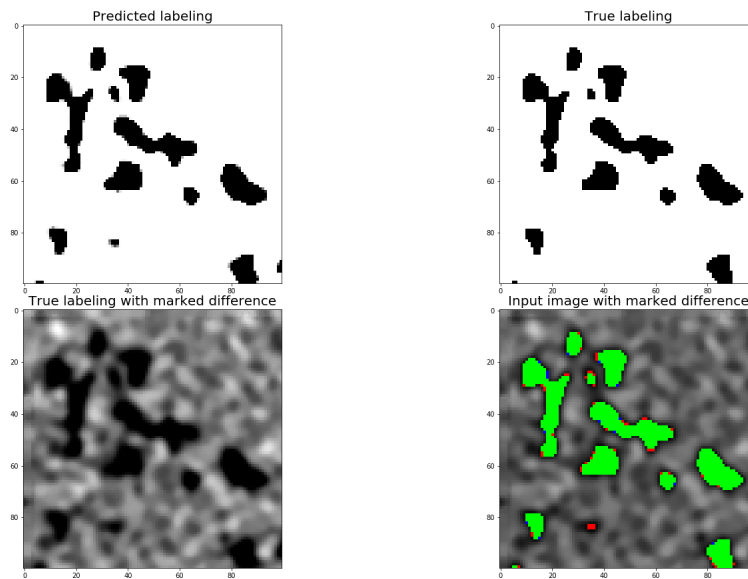


Рис. 5: Результаты модели из группы 3 на образце carb71.

### 6.4 Оценка качества моделей

Для оценки качества моделей использовались следующие метрики:  $\text{logloss}$  (12),  $\text{IOU}$  (13),  $\text{precision}$  – точность предсказания (отношение истинно предсказанных меток первого класса к общему числу предсказанных меток первого класса),  $\text{recall}$  –

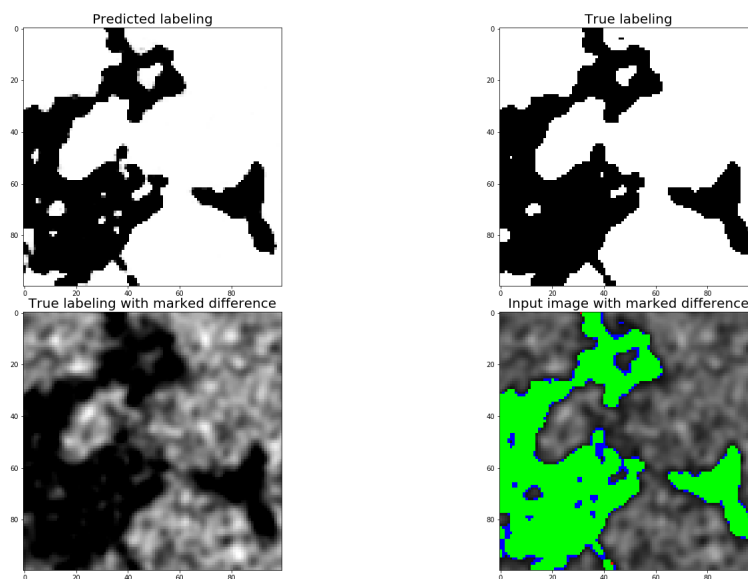


Рис. 6: Результаты модели из группы 2, обученной на образцах Urna-22 и SoilB-2, на образце SPE-carb.

полнота предсказания (отношение истинно предсказанных меток первого класса к общему числу правильных меток первого класса), PR-AUC – площадь под кривой зависимости точности от полноты, ассигасу – верность (число правильно предсказанных меток первого и второго класса к общему числу пикселей).

Результаты для обученных меделей приведены в таблицах 1-8. Таблицы организованны следующим образом: каждая таблица показывает результат работы всех моделей на данном образце, выделенные значения – лучший результат конкретной метрики среди всех моделей.

Результаты можно охарактеризовать следующим образом:

- Модели из группы 1 хорошо работают для образцов того типа, на котором были обучены, на остальных образцах результаты заметно хуже.
- Модели из групп 2, 3 работают хорошо на большинстве образцов, но хуже, чем модели из 1 на специфичных для них образцах.
- Образцы в почвах похожи на песчаники, поэтому модели, обученные на песчаниках, хорошо отрабатывают для почв (в обратную сторону связь не действует, песчаники более специфичной структуры)

- Есть специфичные образцы, на которых все модели работают не лучшим образом(carbRNF), они специально были добавлены для проверки результатов. На этом образце лучше всего работает модель из группы 1 для песчаников, что объясняет специфичность данного образца(карбонат, но похож на песчаник).
- Модель из группы 3 показывает себя относительно хорошо на всех образцах.

Модель	Метрика					
	logloss	PR-AUC	iou	accuracy	precision	recall
carb96558	0,0626	0,9999	<b>0,9883</b>	<b>0,9897</b>	0,9948	0,9935
SoilB-2	0,0856	0,9993	0,9719	0,9747	0,9768	0,9949
Urna-22	0,2514	0,9992	0,9740	0,9771	<b>0,9996</b>	0,9744
Carb96558 + SoilB-2	<b>0,0421</b>	<b>0,9999</b>	0,9882	0,9895	0,9917	<b>0,9965</b>
Carb96558 + Urna-22	0,1248	0,9997	0,9847	0,9865	0,9977	0,9870
SoilB-2 + Urna-22	0,0982	0,9998	0,9863	0,9879	0,9977	0,9886
All	0,0842	0,9999	0,9872	0,9887	0,9974	0,9897

Таблица 1: Результаты для образца carb96558.

Модель	Метрика					
	logloss	PR-AUC	iou	accuracy	precision	recall
carb96558	0,2977	0,9907	0,9317	0,9361	0,9336	<b>0,9978</b>
SoilB-2	1,2784	0,9030	0,8311	0,8346	0,8848	0,9320
Urna-22	<b>0,1345</b>	<b>0,9994</b>	<b>0,9727</b>	<b>0,9760</b>	<b>0,9921</b>	0,9803
Carb96558 + SoilB-2	0,6819	0,9274	0,8992	0,9033	0,9098	0,9872
Carb96558 + Urna-22	0,1551	0,9968	0,9598	0,9637	0,9664	0,9929
SoilB-2 + Urna-22	0,3964	0,9935	0,9223	0,9292	0,9568	0,9624
All	0,2835	0,9890	0,9380	0,9429	0,9474	0,9895

Таблица 2: Результаты для образца carbRNF.

Модель	Метрика					
	logloss	PR-AUC	iou	accuracy	precision	recall
carb96558	0,0473	<b>0,9999</b>	0,9848	0,9861	0,9848	1,0000
SoilB-2	0,0695	0,9998	0,9789	0,9806	0,9790	0,9999
Urna-22	0,0668	0,9999	<b>0,9900</b>	<b>0,9910</b>	<b>0,9966</b>	0,9934
Carb96558 + SoilB-2	0,0590	0,9999	0,9827	0,9841	0,9827	<b>1,0000</b>
Carb96558 + Urna-22	0,0500	0,9990	0,9879	0,9890	0,9883	0,9996
SoilB-2 + Urna-22	0,0379	0,9998	0,9874	0,9886	0,9876	0,9999
All	<b>0,0353</b>	0,9998	0,9869	0,9881	0,9870	0,9999

Таблица 3: Результаты для образца SPE-carb.

Модель	Метрика					
	logloss	PR-AUC	iou	accuracy	precision	recall
carb96558	0,0666	0,9999	0,9886	0,9898	0,9980	0,9906
SoilB-2	0,0349	0,9999	0,9873	0,9885	0,9904	0,9968
Urna-22	0,4208	0,9994	0,9558	0,9604	<b>1,0000</b>	0,9558
Carb96558 + SoilB-2	<b>0,0342</b>	0,9999	0,9884	0,9895	0,9914	<b>0,9969</b>
Carb96558 + Urna-22	0,1753	0,9999	0,9789	0,9811	0,9998	0,9790
SoilB-2 + Urna-22	0,0877	<b>0,9999</b>	0,9861	0,9875	0,9993	0,9868
All	0,0551	0,9999	<b>0,9892</b>	<b>0,9903</b>	0,9978	0,9914

Таблица 4: Результаты для образца SoilAh-1.

Модель	Метрика					
	logloss	PR-AUC	iou	accuracy	precision	recall
carb96558	0,1097	0,9995	0,9762	0,9788	0,9819	0,9941
SoilB-2	0,1414	0,9972	0,9578	0,9615	0,9595	0,9982
Urna-22	0,0728	0,9999	0,9891	0,9904	<b>0,9972</b>	0,9919
Carb96558 + SoilB-2	0,0786	0,9995	0,9788	0,9810	0,9791	0,9996
Carb96558 + Urna-22	0,0276	<b>0,9999</b>	<b>0,9907</b>	<b>0,9917</b>	0,9921	0,9986
SoilB-2 + Urna-22	<b>0,0218</b>	0,9999	0,9904	0,9915	0,9911	0,9993
All	0,0268	0,9999	0,9881	0,9895	0,9884	<b>0,9997</b>

Таблица 5: Результаты для образца TeTree-1

Модель	Метрика					
	logloss	PR-AUC	iou	accuracy	precision	recall
carb96558	0,0828	0,9997	0,9826	0,9845	0,9873	0,9951
SoilB-2	0,1052	0,9986	0,9647	0,9678	0,9660	0,9985
Urna-22	0,0665	0,9999	0,9911	0,9922	<b>0,9985</b>	0,9926
Carb96558 + SoilB-2	0,0521	0,9998	0,9851	0,9867	0,9854	0,9997
Carb96558 + Urna-22	0,0225	<b>1,0000</b>	0,9930	0,9939	0,9945	0,9986
SoilB-2 + Urna-22	<b>0,0165</b>	1,0000	<b>0,9934</b>	<b>0,9942</b>	0,9944	0,9991
All	0,0175	1,0000	0,9917	0,9927	0,9920	<b>0,9997</b>

Таблица 6: Результаты для образца TeTree-2

Модель	Метрика					
	logloss	PR-AUC	iou	accuracy	precision	recall
carb96558	0,2053	0,9988	0,9277	0,9361	0,9301	0,9972
SoilB-2	1,1102	0,8174	0,8219	0,8223	0,8239	0,9970
Urna-22	<b>0,0250</b>	<b>0,9999</b>	<b>0,9868</b>	<b>0,9891</b>	<b>0,9890</b>	0,9978
Carb96558 + SoilB-2	0,3728	0,9883	0,9028	0,9118	0,9055	0,9967
Carb96558 + Urna-22	0,0653	0,9998	0,9733	0,9774	0,9735	<b>0,9997</b>
SoilB-2 + Urna-22	0,0782	0,9995	0,9690	0,9737	0,9707	0,9982
All	0,0971	0,9995	0,9629	0,9684	0,9641	0,9987

Таблица 7: Результаты для образца Urna-30.

Модель	Метрика					
	logloss	PR-AUC	iou	accuracy	precision	recall
<b>carb96558</b>	0,2159	0,9987	0,9306	0,9374	0,9330	0,9973
<b>SoilB-2</b>	1,0337	0,8013	0,8401	0,8401	0,8418	0,9977
<b>Urna-22</b>	<b>0,0509</b>	<b>0,9998</b>	<b>0,9817</b>	<b>0,9843</b>	<b>0,9851</b>	0,9965
<b>Carb96558 + SoilB-2</b>	0,3634	0,9869	0,9057	0,9126	0,9079	0,9974
<b>Carb96558 + Urna-22</b>	0,0708	0,9997	0,9749	0,9784	0,9756	<b>0,9993</b>
<b>SoilB-2 + Urna-22</b>	0,0833	0,9993	0,9716	0,9755	0,9740	0,9975
<b>All</b>	0,1030	0,9992	0,9644	0,9690	0,9653	0,9990

Таблица 8: Результаты для образца Urna-34.



## 7 Заключение

В данной работе построена модель полносвёрточной нейронной сети для сегментации томографических изображений геологических пород.

Представлен алгоритм обучения и применения модели, показана простая возможность перехода от предсказаний вероятностей к бинарным предсказаниям.

Проанализировано качество работы моделей. На основе представленных экспериментальных данных делается вывод, что модели такого вида применимы на практике для широкого спектра пород, а так же возможно использование одной модели для различных типов пород.

В процессе работы обозначилось несколько важных моментов, на которые следует обратить внимание в дальнейших исследованиях:

1. В качестве обучающих данных для моделей в данной работе использовалась сегментация, полученная с помощью метода САС. Несмотря на то, что данный метод настраивал оператор, существуют образцы, качество сегментации которых с помощью данного метода будет низким. Это создаёт проблемы для тренировки любых обучаемых моделей, поэтому данный вопрос требует отдельного исследования.
2. Результаты показывают, что построенные в экспериментах модели, использующие в качестве обучающих данных 2-D изображения, применимы для сегментирования 3-D образцов(секция). Для более естественного учёта 3-D структуры входных данных целесообразно использовать в модели 3-D свёртки, и работать с объёмными данными напрямую.
3. Построение нейросетевых моделей с высокой обобщающей способностью(для многих типов пород) требует более аккуратного и тщательного исследования.

## 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, 2017, V. 179, P. 153-163.
- [2] J. T. Gostick, “Versatile and efficient pore network extraction method using marker-based watershed segmentation” // Physical Review, 2017, E. 96, 023307.
- [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, 2016, V. 7, P. 1125–1139.
- [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, 2016, V. 7, P. 481–492.
- [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, 2016, V. 86, P. 120-128.
- [6] O. Ronneberger , P. Fischer , T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation” // arXiv:1505.04597v1 [cs.CV], 2015.
- [7] L. Mosser , O. Dubrule, , M. J. Blunt, “Reconstruction of three-dimensional porous media using generative adversarial neural networks” // Physical Review, 2017, E. 96, 043309.
- [8] B. L. DeCost , T. Francis , E. A. Holm, “Exploring the microstructure manifold: image texture representations applied to ultrahigh carbon steel microstructures” // arXiv:1702.01117 [cond-mat.mtrl-sci], 2017.
- [9] R. Cang , Y. Xu , S. Chen , Y. Liu , Y. Jiao , M. Y. Ren, “Microstructure Representation and Reconstruction of Heterogeneous Materials via Deep Belief Network for Computational Material Design” // arXiv:1612.07401v3 [cond-mat.mtrl-sci], 2017.

- [10] N. Lubbers , T. Lookman , K. Barros, “Inferring low-dimensional microstructure representations using convolutional neural networks” // Physical Review, 2017, E. 96, 052111.
- [11] A. S. Razavian , H. A. Josephine , S. S. Carlsson, “CNN Features off-the-shelf: an Astounding Baseline for Recognition” // arXiv:1403.6382 [cs.CV], 2014.
- [12] A.P. Sheppard , R.M. Sok, , H. Averdunk, “Techniques for image enhancement and segmentation of tomographic images of porous materials” // Physica A: Statistical Mechanics and Its Applications, 2004, V. 339, P. 145–151.
- [13] H. Deng , D.A. Clausi, “Improved Workflow for Unsupervised Multiphase Image Segmentation” // arXiv:1710.09671 [cs.CV], 2017.
- [14] H. Deng , D.A. Clausi, “Unsupervised image segmentation using a simple MRF model with a new implementation scheme” // Pattern Recognition, 2004, V. 37, P. 2323-2335.
- [15] С.А. Эль-Хатиб, “Сегментация изображений с помощью смешаного и экспоненциального алгоритмов роя частиц” // Информатика и кибернетика, 2015, V. 1.
- [16] M. Pesaresi , J.A. Benediktsson, “A new approach for the morphological segmentation of high-resolution satellite imagery” // IEEE Transactions on Geoscience and Remote Sensing, 2001, V. 39, P. 309-320.
- [17] C. M. Deniz , S. Xiang , S. Hallyburton , A. Welbeck , S. Honig , K. Cho , G. Chang, “Segmentation of the Proximal Femur from MR Images using Deep Convolutional Neural Networks” // arXiv:1704.06176 [cs.CV], 2018.
- [18] D. P. Kingma ,Jimmy Lei Ba, “ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION” // International Conference for Learning Representations, 2015, arXiv:1412.6980 [cs.LG].
- [19] V. Iglovikov , S. Mushinskiy , V. Osin, “Satellite Imagery Feature Detection using Deep Convolutional Neural Network: A Kaggle Competition” // 2017, arXiv:1706.06169v1.