

# ГИБРИДНАЯ МОДЕЛЬ НА ОСНОВЕ U-NET В ЗАДАЧАХ ГЕОЛОГИЧЕСКОЙ ВИЗУАЛИЗАЦИИ

Денисов В.И.<sup>1</sup>, Семенченко Н.Е.<sup>2</sup>

<sup>1</sup> НИ ТПУ, ИШИТР, А2-39, vid3@tpu.ru

<sup>2</sup> НИ ТПУ, ИШИТР, 0В11, nes13@tpu.ru

## Аннотация

В работе представлена улучшенная версия U-Net для сегментации и классификации изображений геологических пород. Модификации включают оптимизацию декодера, альтернативные сверточные слои и механизмы обработки пространственных признаков. Эксперименты показывают, что предложенные изменения повышают качество обработки данных в нефтегазовой отрасли.

**Ключевые слова:** классификация, сегментация, сверточные нейронные сети.

## Введение

Автоматизация процессов с помощью машинного обучения активно проникает во многие промышленные области, включая нефтегазовую и геологическую отрасли. Нейронные сети, в частности, используются для анализа визуальных данных горных пород, получаемых при исследовании керна – образцов, **Ошибка! Источник ссылки не найден.** извлекаемых из скважин (рис. 1).



Рис. 1. Изображение керна

Анализ керна имеет решающее значение для определения свойств пласта и создания гидродинамических моделей месторождений, что является основой для планирования разработки недр [1].

Также исследование керна служит важным инструментом для прогнозирования строения геологического разреза [2–4]. Обзор существующих работ [5–15] показал, что на данный момент нет систем, которые бы обеспечивали достаточно высокую точность и производительность при автоматизированном анализе керна. Имеющиеся подходы к сегментации и классификации образцов пород сталкиваются с рядом ограничений [5–6], что препятствует их широкому внедрению.

Актуальность данной работы определяется потребностью в создании более точной и производительной нейросетевой модели, способной эффективно работать даже при наличии ограниченных или несбалансированных наборов данных. В качестве основы была выбрана архитектура U-Net, показавшая хорошие начальные результаты, но обладающая недостатками в части выразительности извлекаемых признаков, склонности к переобучению и вычислительной сложности.

Для оценки качества разработанного решения использовались стандартные метрики Dice (мера схожести между предсказанной и истинной маской. Чем ближе к 1, тем лучше совпадение) и IoU (отношение площади пересечения к объединению двух масок. Также ближе к 1 – тем лучше). В работе использовался набор высокоразрешённых изображений керна

размером 4000×4000 пикселей, на основе которых вручную создавались маски для обучения модели.

Цель исследования – разработать модифицированную архитектуру на базе U-Net, которая бы демонстрировала повышенную производительность при обработке изображений образцов пород. Оценка эффективности предложенной модификации проводилась с использованием стандартных метрик Dice и IoU.

#### *Модификации сети U-Net*

##### *Функция потерь*

Для оптимизации модели была выбрана бинарная кросс-энтропия (BCE) в качестве функции потерь:

$$L_{BCE} = -\frac{1}{N} * \sum_i [y_i * \log(\sigma(x_i)) + (1 - y_i) * \log(1 - \sigma(x_i))],$$

где  $\sigma$  – сигмоида,  $x_i$  – логит предсказания,  $y_i$  – истинная метка пикселя.

Этот выбор продиктован устойчивостью BCE к дисбалансу классов в данных и ее понятной интерпретацией.

##### *Остаточные блоки*

Проблема затухания градиента в глубоких нейронных сетях, усложняющая обучение ранних слоев, решается путем интеграции остаточных связей, как это предложено в архитектуре ResNet. Выход такого блока рассчитывается по формуле:

$$Y = F(X) + X,$$

где  $X$  – входные данные блока,  $F(X)$  – результат последовательности нелинейных операций (сверток, функций активации, нормализации). Эта техника обеспечивает более свободный проход градиента через сеть, что упрощает обучение и улучшает сходимость. Внедрение остаточных блоков в нашу модель повысило стабильность обучения и качество итоговой сегментации. Прямые «пробросы» внутри блоков предотвращают исчезновение градиента при увеличении глубины сети.

##### *Глубинно-разделимая свёртка*

Вместо стандартных 2D-сверток, которые обрабатывают пространственные и каналные измерения одновременно, были использованы глубинно-разделимые свертки. Они разделяют операцию на два этапа:

- *Depthwise Convolution*: Применение отдельных фильтров к каждому входному каналу.
- *Pointwise Convolution (1x1)*: Линейное комбинирование выходных каналов предыдущего этапа.

Данный подход значительно уменьшает количество обучаемых параметров и вычислительную сложность модели, практически не влияя на ее точность. Это ускоряет как обучение, так и процесс инференса.

##### *Промежуточные связи между слоями*

Для более эффективного объединения признаков с разных уровней иерархии, были добавлены дополнительные промежуточные связи между слоями декодера. Эта модификация, вдохновленная архитектурой U-Net++ [16-17], способствует созданию более насыщенных и плавных карт признаков. Такая структура особенно полезна при анализе геологических пород, где необходимо различать тонкие текстурные и структурные особенности разных литотипов, так как улучшается точность сегментации объектов со сложной морфологией.

##### *«Онлайн» аугментация*

Аугментация данных – стандартный прием для повышения робастности моделей и предотвращения переобучения путем искусственного расширения датасета. Классический («оффлайн») подход требует предварительной генерации и сохранения большого объема измененных изображений. Это затратно с точки зрения дискового пространства и ограничивает разнообразие трансформаций. В данной работе реализован метод «онлайн» аугментации, выполняемой непосредственно во время обучения. Алгоритм представлен на блок-схеме (рис. 2).

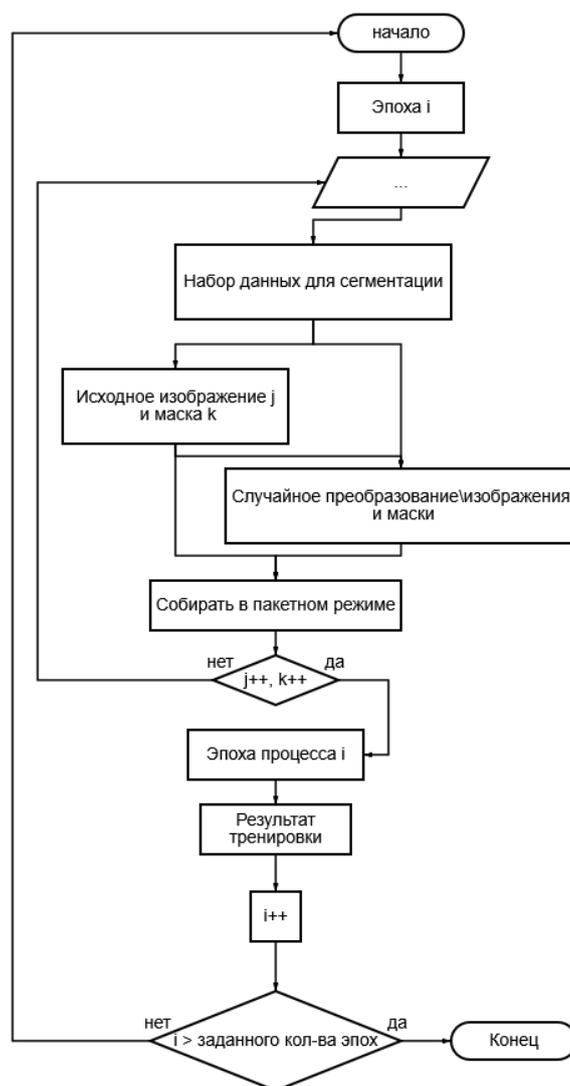


Рис. 2. Блок-схема разработанного алгоритма аугментации

При каждой загрузке батча данных в эпоху обучения, для части изображений генерируются случайные трансформации (например, повороты, сдвиги, изменение яркости). Таким образом, модель на каждой итерации видит немного измененные версии одних и тех же данных, что эквивалентно обучению на значительно большем и более разнообразном датасете без необходимости его физического хранения. Этот подход позволил создать модель, более устойчивую к вариациям входных данных и обладающую лучшей обобщающей способностью.

### Результаты модификации сети

Для оценки эффективности предложенной усовершенствованной архитектуры U-Net было проведено сравнение с базовой U-Net. Ключевые архитектурные отличия см. в таблице 1.

Таблица 1. Сравнение ключевых характеристик архитектур

Характеристика	Оригинальная U-Net	Модифицированная U-Net
Блоки сверток	Классические сверточные блоки	Остаточные блоки с пропусками
Тип сверток	Стандартные (2D) свертки	Глубинно-разделяемые свертки
Механизмы внимания	Отсутствуют	AttentionBlock для выделения значимых областей
Декодер	Простая конкатенация и свертки	Декодер с механизмами внимания и остаточными соединениями

Сравнение проводилось на задаче сегментации изображений образцов керна, полученных из лабораторных исследований. Был подготовлен соответствующий датасет изображений и масок. Для валидации и финальной оценки использовалась отложенная тестовая выборка, составляющая 20% от всего набора данных. Сравнение оригинальной и модифицированной нейронной сети показано на рисунке 3.

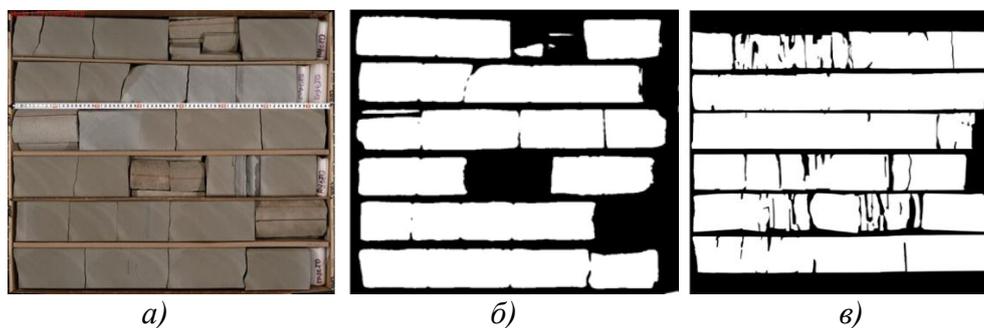


Рис. 3. а) исходное изображение б) пример работы U-Net в) пример работы модифицированной U-Net

Сравнение характеристик архитектур приведено в таблице 2.

Таблица 2. Сравнительная характеристика архитектур сетей

Модель	Train Dice	Train IoU	Val Dice	Val IoU	Время обучения (25 эпох)	Среднее время обработки 1 изображения	Используемая память (ГБ)
U-Net	0,9220	0,8870	0,9060	0,8540	~2 часа	~0,15 сек	~4,0 ГБ
Модифицированная U-Net	0,9807	0,9625	0,9667	0,9361	~1,8 часа	~0,10 сек	~7,0 ГБ

Предложенная модифицированная U-Net продемонстрировала существенное превосходство над стандартной архитектурой U-Net по всем ключевым показателям качества. На валидационной выборке значение метрики Dice возросло на 6.07 % (с 0.9060 до 0.9667), а метрики IoU – на 8.71 % (с 0.8540 до 0.9361). Помимо увеличения точности и качества сегментации, модифицированная архитектура обеспечивает улучшение производительности модели: среднее время обработки одного изображения сократилось на 33 % (с 0.15 сек до 0.10 сек), а время обучения на 25 эпох – на 10 % (с ~2 до ~1,8 часа).

Однако стоит отметить, что модифицированная U-Net++ требует большего объема оперативной памяти – на 75 % выше (7.0 ГБ против 4.0 ГБ у стандартной U-Net), что является приемлемым компромиссом в контексте значительно улучшенных показателей качества и скорости работы модели.

Внедрение «онлайн»-аугментации в модифицированную модель дополнительно способствовало улучшению её обобщающей способности, снизило вероятность переобучения и повысило стабильность результатов.

Таким образом, модифицированная архитектура U-Net++ с онлайн-аугментацией является оптимальным и эффективным решением для задач сегментации образцов керна, обеспечивая превосходство по точности и производительности над базовой U-Net.

## Заключение

В работе предложена усовершенствованная архитектура на базе U-Net, адаптированная для задач классификации и сегментации образцов геологических пород. Основные внесенные изменения включают:

- Использование остаточных блоков (ResidualConvBlock) для улучшения прохождения градиента и стабилизации обучения.
- Применение глубинно-разделимых сверток для уменьшения числа параметров и ускорения вычислений.
- Внедрение промежуточных связей в декодере (по типу U-Net++) для более точной реконструкции деталей и сегментации мелких объектов.
- Интеграция блоков внимания (Attention Blocks) для фокусировки модели на наиболее релевантных участках изображения.

Кроме того, использование “онлайн” аугментации позволило повысить робастность модели и эффективность использования обучающих данных. Экспериментальные результаты подтверждают, что предложенная модель превосходит как классическую U-Net, так и U-Net++ по метрикам качества (Dice, IoU), демонстрируя более быструю сходимость и повышенную точность сегментации, особенно на сложных участках изображений.

## Список использованных источников

1. Гурбатова И.П., Плотников В.В., Попов Н.А. Особенности изучения фильтрационных характеристик ориентированного зерна сложнопостроенных карбонатных коллекторов // Вестник ПНИПУ. Геология. Нефтегазовое и горное дело. – 2013. – № 9. – С. 79-86.
2. Иванченков В.П., Кочегуров А.И., Нгуен С.Х. Фазочастотный алгоритм прослеживания сейсмических сигналов с управляемой протяженностью функции качества // Научный вестник Новосибирского государственного технического университета. – 2014. – № 4. – С. 59-68.
3. Иванченков В.П., Кочегуров А.И., Орлов О.В. Применение методов фазочастотного анализа сейсмических волн для изучения околоскважинного пространства // Нефтяное хозяйство. – 2015. – № 10. – С. 58–63.
4. Кочегуров А.И., Иванченков В.П., Орлов О. В. Прогноз геологического разреза на основе детальной обработки сейсмических материалов в программно-алгоритмическом комплексе «ГЕОСЕЙФ» // Известия томского политехнического университета. Инжиниринг георесурсов. – 2019. – Т. 330, № 6. – С. 134-144.
5. Кочегуров А.И., Денисов В.И., Задорожных Е.А. Анализ применения методов машинного обучения в задачах классификации пород на образцах зерна // Известия Томского политехнического университета. Инжиниринг георесурсов. – 2024. – Т. 335. – № 9. – С. 148-159.
6. Денисов В.И., Задорожных Е.А. Применение средств машинного обучения для исследования образцов зерна // Сборник трудов XXI Международной научно-практической конференции студентов, аспирантов и молодых ученых МСИТ-2024. – Томск : изд-во ТПУ. – 2024. – С. 155-160.
7. Patel A.K., Snehmoy C. Computer vision-based limestone rock-type classification using probabilistic neural network // Geoscience Frontiers. – 2016. – № 7. – P. 53-60.
8. Gorai A.G. and others. Design and development of a machine vision system using artificial neural network-based algorithm for automated coal characterization // International Journal of Coal Science & Technology. – 2020. – №8. – P. 737-755.
9. Patel A., Snehmoy C., Gorai A.K. Development of online machine vision system using support vector regression (SVR) algorithm for grade prediction of iron ores // Fifteenth IAPR International Conference on Machine Vision Applications (MVA). – Nagoya, 2017. – P. 149-152.
10. Ивченко А.В., Барабошкин Е.Е., Исмаилова Л.С. Литологическая интерпретация фотографий зерна на основе компьютерного анализа // Международная научная конференция

«IEEE Northwest Russia Conference On Mathematical Methods In Engineering And Technology: ММЕТ NW 2018». – Санкт-Петербург : СПбГЭТУ «ЛЭТИ», 2018. – С. 430-433.

11. Артамонов В.П., Каюров Н.К. Автоматизация алгоритмов предобработки и анализа фотографий керн для целей процесса макро- и литологического описания // III научно-практическая конференция «Интеллектуальный анализ данных в нефтегазовой области». – Новосибирск, 2022. – С. 14-16.

12. Дьячков С.М. Автоматическое определение породы по фотографиям керн современными методами машинного обучения // Вестник Тюменского государственного университета. Серия: Физико-математическое моделирование. Нефть, газ, энергетика. – Тюмень: Издательство Тюменского государственного университета. – 2021. – Т. 7, № 4(28). – С. 181-198.

13. Постников А.В., Постникова О.В., Хасанов И.И. Новые методические аспекты литологических исследований пород бажендовской свиты // Нефтяное хозяйство. – 2015. – № 10. – С. 23-27.

14. Хасанов И.И. Анализ цветности пород по цифровым изображениям керн // Геология нефти и газа. – 2014. – № 5. – С. 33-39.

15. Khasanov I. The use of computer analysis of digital images for the study of core material // Geomodel 2015 - 17th science and applied research conference on oil and gas geological exploration and development. – 2015. – V. 2015. – P. 1-5.

16. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation // MICCAI 2015. Lecture Notes in Computer Science. – 2015.

17. Zongwei Zhou, Jianming Liang, Nima Tajbakhsh. UNet++: A Nested U-Net Architecture for Medical Image Segmentation // Conference: 4th Deep Learning in Medical Image Analysis (DLMIA) Workshop. – Spain. – 2018.