

Анализ влияния использования кроссполяризованной съемки и предобработки микроизображений аншлифов на нейросетевую сегментацию минералов

Краснова С. М.^{1*}, Хвостиков А. В.¹, Коршунов Д. М.²,
Крылов А. С.¹, Сорокин Д. В.^{1*}

¹Факультет вычислительной математики и кибернетики,
Московский государственный университет, Москва, Российская
Федерация.

²Геологический институт, Российская академия наук, Москва,
Российская Федерация.

*Автор(ы), ответственный(ые) за переписку. E-mail(s):
skrasnova005@gmail.com; dsorokin@cs.msu.ru;

Соавторы: khvostikov@cs.msu.ru; d.korshunov@ginras.ru;
kryl@cs.msu.ru;

Аннотация

Задача автоматической сегментации оптических микроскопических изображений аншлифов является актуальной для решения проблем геологии и минералогии. Особую сложность представляет анализ анизотропных минералов, оптические свойства которых изменяются в зависимости от ориентации в кросс-поляризованном свете. В данной работе исследуется гипотеза о том, что использование серии поворотных изображений в кросс-поляризованном свете может повысить точность их автоматического распознавания. Проводится расширенное и строгое исследование этой гипотезы на полном наборе данных LumenStone, включающем три различные минеральные ассоциации (S1, S2, S3). Для совмещения изображений применён алгоритм LoFTR, проведены многократные запуски обучения для оценки устойчивости, проанализировано влияние предобработки изображений в виде выравнивания освещенности. Показано, что добавление поворотных изображений не приводит к статистически значимому улучшению средней точности сегментации по сравнению с базовой моделью. Однако выравнивание освещенности демонстрирует стабильный положительный эффект, существенно снижая вариативность результатов. Наиболее эффективной и стабильной конфигурацией из

рассмотренных является базовая архитектура ResUNet, примененная к данным после корректировки освещенности.

Ключевые слова: Сегментация изображений, Совмещение изображений, Геология, Нейронные сети, Аншлафы геологических образцов, Оптическая микроскопия, Анизотропные минералы

Получено редакцией 28.12.2025; внесены авторские правки 30.12.2025; принята к публикации 30.12.2025

1 Введение

В современной геологии важной задачей является анализ минерального состава руд и пород [1, 2]. Этот анализ включает определение минералов, их соотношение и структурные особенности.

Анализ минерального состава руд важен для поиска месторождений и выбора методов добычи полезных ископаемых. Также он необходим для контроля качества сырья на промышленных предприятиях, чтобы обеспечить эффективное производство.

Для определения состава минералов используются разные методы, такие как химический [3] и физико-химический анализ. Эти методы точны, но требуют времени, денежных средств и усилий. Существуют также оптические методы [4, 5], которые позволяют изучать образцы руд под микроскопом. Эти методы дешевле, но требуют высокой квалификации геологов и все равно могут быть сложными и трудоемкими. Эти вызовы стимулируют развитие и внедрение алгоритмических методов сегментации, которые используют только цифровое изображение образца как исходные данные. Развитие в этой области идет от классических методов обработки изображений к современным, основанным на глубоком обучении, что позволяет значительно повысить скорость, точность и воспроизводимость анализа структур горных пород.

Для решения задачи автоматической сегментации геологических изображений могут применяться классические методы компьютерного зрения. Один из них заключается в предварительном выделении суперпикселей [6–8] (небольшие кластеры пикселей, объединенные по определенным признакам). Каждому кластеру сопоставляется вектор признаков, который описывает характеристики кластера, например, цветовые компоненты или текстурные особенности. На основании этих векторов производится дальнейшая классификация, сегментирующая суперпиксели в заданные классы. Такой подход снижает вычислительную нагрузку, так как вместо обработки каждого пикселя по отдельности анализируются суперпиксели, что уменьшает объем вычислений. Существуют разные методы выделения суперпикселей, такие как, например,

SLIC [9, 10] (Simple Linear Iterative Clustering) и Watershed Segmentation [11].

Однако методы, основанные на суперпикселях, часто имеют ограничения при работе со сложными текстурами аншлифов, где вариативность оптических свойств (цвета, яркости, текстуры) внутри одного класса может быть высока. Также эффективность таких методов сильно зависит от ручного выбора параметров. Вследствие чего, в последние годы нейросетевые методы сегментации достигли значительных успехов в обработке изображений, в том числе в области анализа геологических данных [12]. Один из наиболее эффективных методов – это использование нейросетевых архитектур, таких как U-Net, специально разработанных для сегментации изображений [13]. Эта архитектура также успешно применяется для сегментации геологических изображений [12].

Предобработка изображений аншлифов геологических образцов (полированных образцов руд) играет ключевую роль в обеспечении точности и надежности последующего анализа. В данной работе уделяется внимание такому виду предобработки, как выравнивание освещенности.

Искажения освещенности возникают из-за неравномерного распределения света на поверхности образца. Это может быть вызвано различными факторами, такими как неправильное расположение источника света, отражения от поверхности образца или неравномерное поглощение света различными минералами. Такие искажения могут привести к неправильной интерпретации данных, так как различные участки изображения будут иметь разное освещение, что затрудняет точную идентификацию минералов. Выравнивание освещенности позволяет устранить эти проблемы, обеспечивая равномерное освещение, что значительно улучшает качество изображения и облегчает последующий анализ минералов.

Для выравнивания освещенности на изображениях разработано множество методов. Одним из известных направлений являются алгоритмы, использующие гистограммы для корректировки интенсивности освещения [14, 15]. Эти подходы предполагают анализ распределения интенсивностей пикселей изображения. Другим популярным подходом, хорошо устраняющим неравномерности освещения, сохраняя детали сцены, является модель Retinex [16, 17], основанная на разделении изображения на компоненты отражения и освещения:

$$I(x, y) = R(x, y)L(x, y), \quad (1)$$

где $I(x, y)$ – наблюдаемое изображение, $R(x, y)$ – отражающая способность объекта, а $L(x, y)$ – карта освещения. Основная цель – восстановить функцию $R(x, y)$, которая характеризует истинные свойства сцены.

Анизотропные минералы характерны тем, что изменяют свои оптические свойства в зависимости от ориентации относительно направления поляризованного света. В условиях кросс-поляризованной съёмки, при

вращении предметного столика микроскопа, такие минералы демонстрируют эффект периодического изменения цвета и яркости (см. Рис. 1). Это явление обусловлено особенностями их кристаллической структуры.

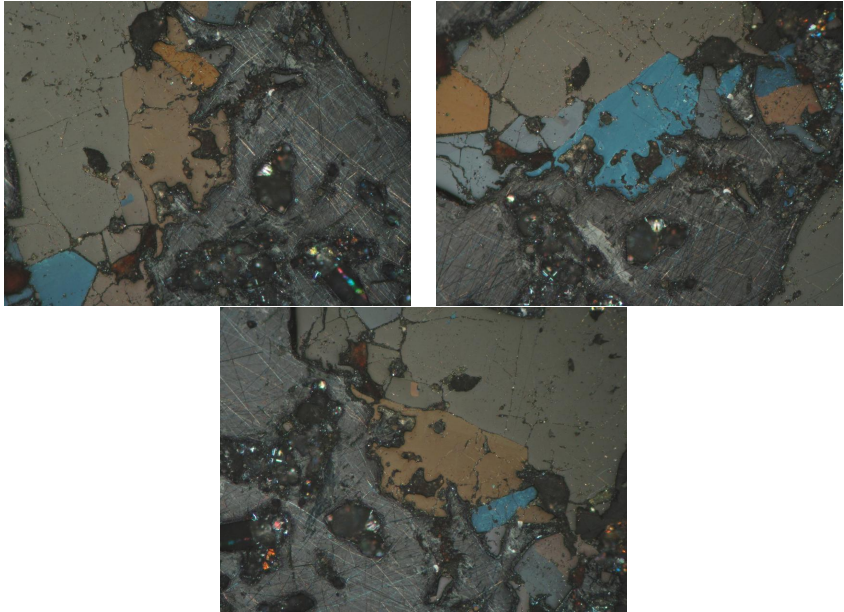


Рис. 1: Микроскопические изображения аншлифа, содержащего анизотропные минералы, под разными ориентациями: 0, 60 и 120 градусов.

В ряде случаев два разных минерала могут иметь схожий цвет и текстуру на изображении, полученном в плоско-поляризованном свете, и становятся различимыми только за счёт анализа их поведения в кросс-поляризованном свете при изменении угла наблюдения.

Настоящее исследование представляет собой углубленный анализ подхода, впервые предложенного в работе [18]. В указанной работе было выдвинуто предположение, что использование серии изображений в кросс-поляризованном свете может повысить точность сегментации анизотропных минералов. Данная статья проводит более широкое и строгое исследование этой гипотезы, внося следующие ключевые усовершенствования в экспериментальную базу и методику:

1. Усовершенствование метода совмещения изображений: замена алгоритма SIFT [19] на более современный и точный LoFTR [20] для повышения точности совмещения поворотных изображений.

2. Проведение многократных запусков обучения для статистически значимой оценки метрик.

3. Использование расширенного набора данных для более репрезентативных выводов.

4. Исследование влияния предобработки: анализ этапа выравнивания освещенности.

Проведенный комплексный анализ позволяет сделать более взвешенные выводы о влиянии дополнительных поворотных изображений. В частности, установлено, что их использование не приводит к значительному статистическому улучшению метрик по сравнению с базовым подходом (не использующим дополнительные изображения), однако выравнивание освещенности демонстрирует стабильный положительный эффект на воспроизводимость результатов.

2 Набор данных

Для обучения и оценки моделей в работе используется общедоступный набор данных LumenStone¹, специально созданный для задач анализа геологических изображений. Этот набор объединяет высококачественные изображения аншлифов различных минеральных ассоциаций с предоставленными пиксельными масками. Для задачи сегментации предоставлены три независимых поднабора (S1, S2, S3), каждая часть представляет собой отдельное месторождение или тип руды, что обеспечивает разнообразие данных для проверки обобщающей способности моделей.

- S1 содержит изображения гидротермальных руд Берёзовского месторождения (сфалерит, пирит, галенит, борнит, тетраэдрит-теннантит, халькопирит).
- S2 включает образцы расслоенных ультрамафитовых месторождений Норильской группы (пирротин, халькопирит, пентландит, магнетит).
- S3 представляет высокотемпературные гидротермальные руды (пирит, арсенопирит, ковеллин, борнит, халькопирит, магнетит, гематит).

В предыдущих исследованиях, посвященных сегментации аншлифов с использованием набора данных LumenStone, применялись лишь отдельные его части. Например, в работе [12] обучение и оценка проводились только на наборе S1, а в исследовании [18] – на S1 и S3. В данной работе представлен комплексный анализ качества сегментации на наборах LumenStone S1, S2 и S3, что позволяет получить более репрезентативные и обобщённые выводы о производительности методов.

Важной особенностью, используемой в данной работе, является наличие поворотных серий для части изображений в наборах S2 (3 изображения) и S3 (22 изображения). Поворотная серия – это набор снимков одного и того же участка аншлифа в кросс-поляризованном свете,

¹<https://imaging.cs.msu.ru/en/research/geology/lumenstone>

сделанных с шагом вращения столика микроскопа в 15° . Эти данные позволяют исследовать влияние информации об анизотропных оптических свойствах минералов на качество их автоматической сегментации.

3 Методы

3.1 Совмещение изображений

Поскольку используемые в работе серии изображений в кросс-поляризованном свете получены при разных углах поворота столика микроскопа, необходимо обеспечить точное пространственное совмещение изображений. Для этого применяется LoFTR (Local Feature Transformer) [20] – метод сопоставления ключевых точек на основе трансформеров, устойчивый к изменениям освещения и текстуры.

LoFTR анализирует изображения, выделяя устойчивые локальные особенности. Найденные соответствия между изображениями позволяют вычислить аффинное преобразование для точного совмещения.

Все дополнительные изображения под кросс-поляризованным светом приводятся к единой системе координат относительно основного снимка под плоско-поляризованным светом (см. Рис. 2).

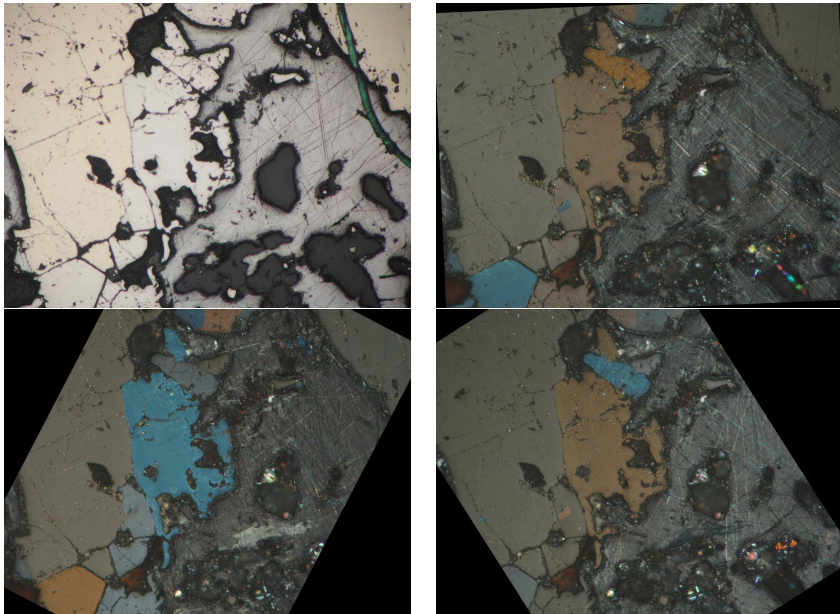


Рис. 2: Первое изображение – основной снимок под плоско-поляризованным светом, затем 3 дополнительных изображения под кросс-поляризованным светом, совмещенных с основным изображением.

3.2 Выравнивание освещенности

Для обеспечения универсальности применения процедуры предобработки изображений в условиях, характерных для различных типов микроскопов, предполагается создание метода, способного автоматически выравнивать освещенность на изображениях шлифов. Основное требование к методу заключается в его доступности и простоте использования: любой специалист-геолог, независимо от уровня подготовки и особенностей используемого оборудования, должен иметь возможность получить скорректированное изображение с минимальными усилиями.

Функционал метода разработан с учетом разнообразия моделей микроскопов и параметров их оптических систем. Программное обеспечение должно обеспечивать обработку изображений, полученных на различных устройствах, автоматически адаптируясь к индивидуальным характеристикам каждого микроскопа. Такой подход исключает необходимость предварительной настройки системы и позволяет геологу сосредоточиться непосредственно на исследовательских задачах.

В данной работе для выравнивания освещенности используется концепция мультипликативной модели, вдохновленной принципами модели Retinex [16, 17], описываемой уравнением 1. Однако предложенная в работе модель была адаптирована и модифицирована в контексте задачи выравнивания освещенности на геологических изображениях. Было установлено, что для конкретного микроскопа можно выделить отдельную карту освещенности $L(x, y)$, которая будет характерна именно для этого устройства. Такой подход значительно ускоряет процесс обработки изображений, поскольку позволяет применять одну и ту же карту освещенности ко всем изображениям, получаемым при исследовании на данном микроскопе, что делает метод более эффективным и удобным для практического использования геологами

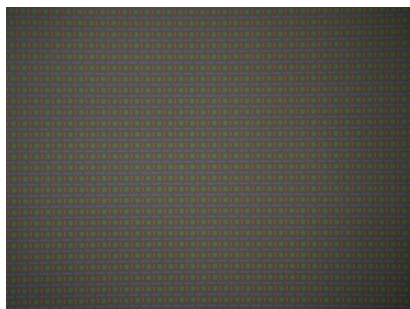


Рис. 3: Используемая фотография матрицы дисплея смартфона под микроскопом.

Предложенный в данной работе метод выравнивания освещенности основан на анализе фотографии дисплея телефона с OLED матрицей, сделанной под микроскопом (см. Рис. 3). На таких изображениях хорошо различимы красные, синие и зеленые пиксели, характерные для структуры дисплея. Алгоритм работы включает следующие этапы:

Выделение одного канала изображения. На первоначальном этапе выбирается один цветовой канал. Алгоритм анализирует центральную часть изображения (40-60% от его ширины и высоты) и выбирает в качестве основного тот цветовой канал (красный, зелёный или синий), который является наиболее тёмным в этой области. Предполагается, что именно в этом канале контраст между пикселями и фоном выражен сильнее всего, что делает его наиболее информативным для дальнейшей обработки. Однако после такого выделения на изображении присутствуют пиксели из других каналов.

Локальный пороговый анализ. Для устранения нежелательных пикселей проводится обработка изображения с использованием локального порогового значения яркости с учётом дисперсии. Изображение делится на окна размером 128 x 128. Для каждого окна вычисляются среднее значение яркости и стандартное отклонение. Порог устанавливается как сумма среднего значения и 1.5-кратного стандартного отклонения. Пиксели, чья яркость выше порога, считаются принадлежащими выбранному на 1 шаге каналу дисплея и получают значение 1 (белый цвет) в бинарном изображении. В каждой локальной области остаются только самые яркие пиксели, что позволяет почти идеально выделить точки, принадлежащие выбранному каналу дисплея.

Фильтрация шума. На этом этапе применяется медианная фильтрация для сглаживания изображения, а также фильтрация выделенных точек по их площади. Сначала изображение обрабатывается с помощью медианного фильтра с ядром 9 x 9. Затем находятся контуры имеющихся на изображении объектов, и для каждого вычисляется площадь. Адаптивный порог для площади устанавливается как разность среднего значения площадей и 1.5-кратного стандартного отклонения. Все точки, площадь которых меньше адаптивного порога, удаляются. Это позволяет исключить шумовые артефакты и сохранить только значимые пиксели.

Определение центров точек и средней интенсивности. После выделения точек для выбранного канала находятся их центроиды, а также вычисляется средняя интенсивность для каждой точки.

Аппроксимация функцией Гаусса. Центроиды точек и их интенсивности используются для построения двумерной карты освещенности, которая аппроксимируется функцией Гаусса:

$$G(x, y) = Ae^{-\left(\frac{(x-x_0)^2}{2\sigma_x^2} + \frac{(y-y_0)^2}{2\sigma_y^2}\right)}, \quad (2)$$

где множитель A используется для масштабирования амплитуды гауссовой функции, чтобы она соответствовала диапазону значений интенсивности на дискретной карте освещенности. Параметры функции $(A, x_0, y_0, \sigma_x, \sigma_y)$ определяются методом наименьших квадратов с использованием алгоритма Левенберга–Марквардта. Этот шаг позволяет сгладить карту освещенности и устранить мелкие локальные отклонения. Для применения модели коррекции $R(x, y) = \frac{I(x, y)}{L(x, y)}$ необходимо, чтобы центральная (наиболее яркая) область карты $L(x, y)$ имела значение, близкое к единице, иначе в центре изображения произойдет нежелательное изменение яркости. Поэтому выполняется аддитивное смещение: ко всем значениям карты прибавляется константа δ , вычисляемая как $\delta = 1 - \max_{(x, y)} L(x, y)$. В результате максимальное значение карты становится равным 1, а её форма и относительные различия между пикселями сохраняются (см. Рис. 4). После этого скорректированная карта $L_{\text{норм}}(x, y) = L(x, y) + \delta$ используется для выравнивания освещенности исходного изображения.



Рис. 4: Построенная карта освещенности.

Описанный подход позволяет эффективно компенсировать влияние неоднородного освещения, сохраняя структурные детали изображения. Пример применения метода приведен на Рис. 5

3.3 Балансировка классов

Важной особенностью набора данных является то, что классы несбалансированы. В наборе изображений есть классы, представленные очень маленьким количеством пикселей по сравнению, например, с фоновыми пикселями. Такой дисбаланс данных может привести к снижению способности модели к обучению и негативно повлиять на качество сегментации. Для решения этой проблемы в данной работе используется метод балансировки выбора классов, представленный в [21].

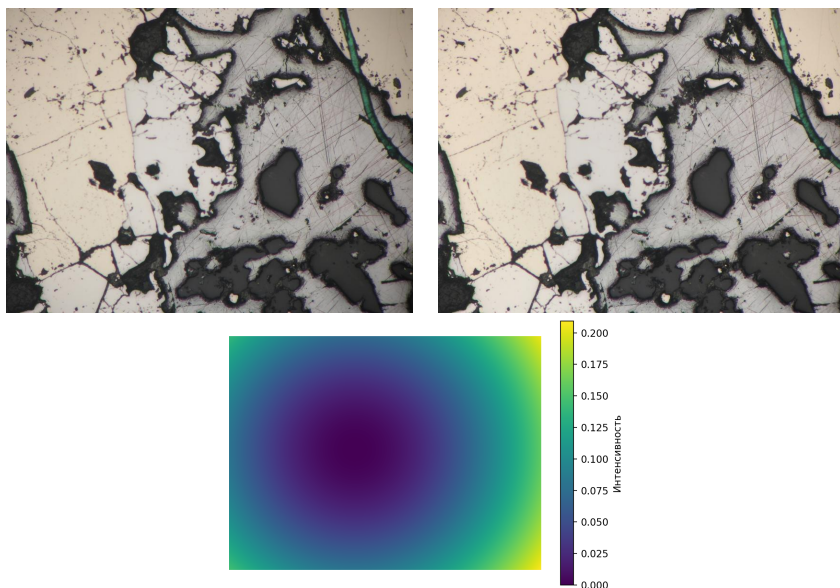


Рис. 5: Первое изображение – исходное, второе – с выровненной освещенностью. Третье изображение – цветная визуализация разностного изображения (дельта-карта) между скорректированным и исходным кадрами с сопутствующей шкалой.

Метод использует трехэтапный отбор данных для обучения модели:

1. На первом этапе определяется, какой именно класс минерала будет представлен в текущем фрагменте. Выбор осуществляется с учётом текущей статистики накопленного числа пикселей по каждому классу: классы, которые пока встречаются реже, имеют более высокую вероятность быть выбраны.

2. После того как выбран целевой класс, из всех изображений, содержащих этот класс, отбирается одно с использованием весов, зависящих от количества пикселей данного класса на каждом изображении.

3. Из выбранного изображения извлекается фрагмент с интересующим классом.

Для каждого изображения из набора алгоритм строит карту вероятности выбора фрагмента. Эта карта определяет, какие области изображения с наибольшей вероятностью будут выбраны в качестве фрагмента для обучения. Чем выше вероятность в данной точке, тем выше шанс, что она станет частью извлекаемого фрагмента.

В данной работе исследовались в том числе анизотропные минералы, образцы которых сканировались под разными углами поляризации. При совмещении таких изображений возникают пустые участки (см. Рис. 2),

не содержащие полезных данных. Чтобы исключить выбор фрагментов из этих областей, исходная карта вероятностей умножается на бинарную маску допустимой зоны, которая выделяет только релевантные области с данными минералов. Таким образом, алгоритм гарантирует, что обучающие фрагменты извлекаются исключительно из информативных частей изображения. Пример допустимой зоны и карт вероятностей, построенных с использованием этого алгоритма, показан на Рис. 6.

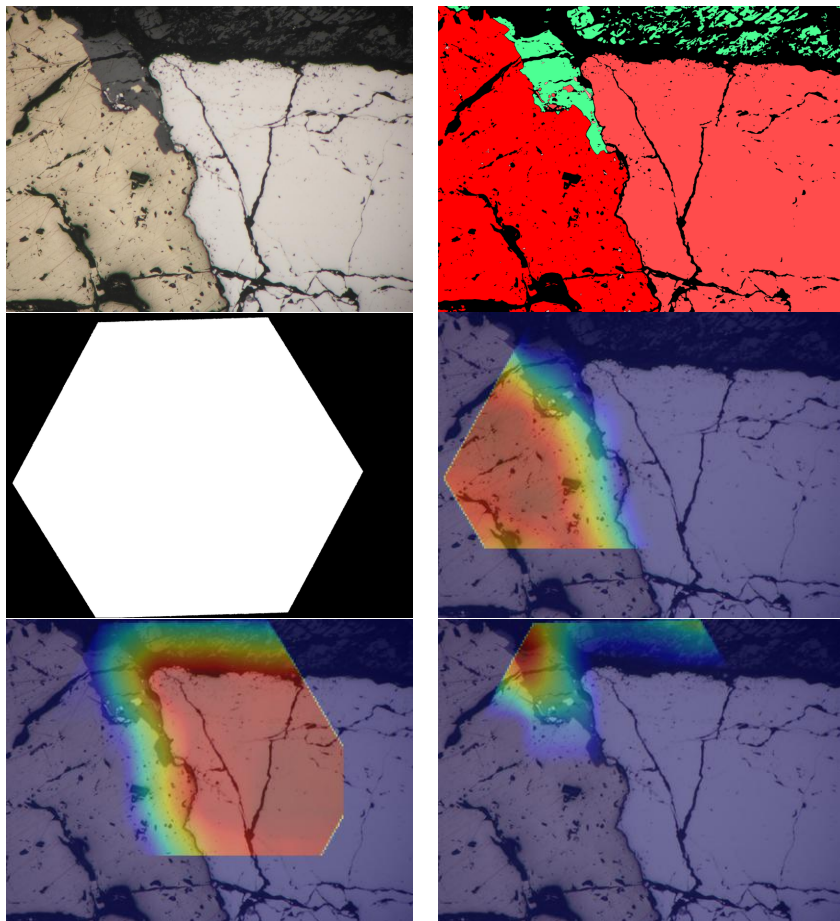


Рис. 6: Первое – исходное изображение, второе – его ручная сегментация, третье – допустимая зона. Затем – карты вероятностей, наложенные на исходное изображение, для классов халькопирит, пирит и сфалерит.

3.4 Архитектура нейронной сети для сегментации

Для классификации минералов применяется нейросетевая архитектура ResUNet [22], объединяющая преимущества остаточных связей (ResNet) [23], UNet [24] и дилатационных сверток.

На вход нейронной сети подается основное изображение под плоско-поляризованным светом, а также 1, 3 или 6 дополнительных совмещенных изображения под кросс-поляризованным светом. Если дополнительные изображения отсутствуют, соответствующие каналы заполняются нулями, что позволяет модели работать даже с неполными данными (см. Рис. 7). Для ускорения обучения исходное разрешение изображений уменьшено в 2 раза. В нейронную сеть подается набор изображений, содержащий 64 фрагмента, каждый размером 384×384 . Таким образом, входной тензор имеет размерность $[64, 384, 384, 3 \cdot N]$, где N – количество дополнительно подаваемых изображений в кросс-поляризованном свете.

Архитектура ResUNet реализованы на Python 3.10 с использованием библиотеки PyTorch. В качестве оптимизатора выбран Adam [25] с начальной скоростью обучения 0.001. Каждая сеть обучалась в течение 50 эпох, каждая эпоха включает 160 итераций. Для оценки качества работы моделей во время обучения в конце каждой эпохи используется 80 наборов из 64 случайных фрагментов размером 384×384 тренировочной выборки данных. В качестве функции потерь используется многоклассовая энтропия.

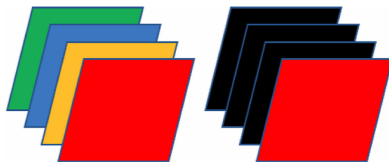


Рис. 7: Вид фрагмента из входного тензора. Слева - вид фрагмента в случае известных дополнительных изображений. Справа - вид фрагмента в отсутствие дополнительных изображений. Красным цветом обозначено основное изображение в плоско-поляризованном свете.

4 Результаты

В данном разделе представлены результаты сравнения качества сегментации для четырёх конфигураций модели:

- Baseline: классическая архитектура ResUNet, использующая только одно основное изображение в плоско-поляризованном свете и не использующая дополнительную информацию.

Таблица 1: Качество сегментации (IoU) на исходных данных. **Жирным шрифтом** выделены анизотропные минералы, представленные в поворотных сериях, а также наилучшее значение IoU для каждой строки (минерала).

Название минерала (аббревиатура)	Baseline	M1	M3	M6
Background (BG)	0.9019	0.9051	0.9024	0.9042
Chalcopyrite (Ccp)	0.8155	0.8068	0.7872	0.7802
Galena (Gl)	0.6236	0.6169	0.578	0.5759
Magnetite (Mag)	0.352	0.4507	0.3974	0.3421
Bornite (Br)	0.7705	0.7923	0.7698	0.7819
Pyrite (Py)	0.7721	0.7557	0.7392	0.7481
Sphalerite (Sph)	0.7402	0.7554	0.7402	0.7478
Tennantite (Tnt)	0.6708	0.6644	0.6508	0.6744
Pentlandite (Pn)	0.5228	0.5373	0.4801	0.4791
Covellite (Cv)	0.7377	0.7162	0.6861	0.6901
Hematite (Hem)	0.0464	0.0955	0.0793	0.0694
Arsenopyrite (Apy)	0.7278	0.7282	0.6887	0.7056
Marcasite (Mrc)	0.41	0.4031	0.3723	0.2714
Pyrrhotite (Po)	0.808	0.7944	0.7848	0.7936
Mean IoU \pm std:	0.6357 \pm 0.0274	0.6444 \pm 0.0179	0.6183 \pm 0.0122	0.6117 \pm 0.0269

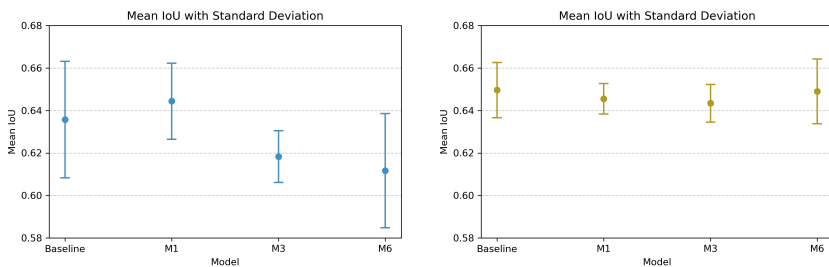


Рис. 8: Среднее значение IoU и стандартное отклонение для четырёх моделей сегментации изображений. Первый график – модели обучались на исходном наборе данных, второй – модели обучались на наборе данных с выровненной освещенностью.

- M1: модель, принимающая на вход основное изображение и одно дополнительное совмещённое изображение того же участка, полученное в кросс-поляризованном свете при 0° вращения.

Таблица 2: Качество сегментации (IoU) на данных с выровненной освещенностью. **Жирным шрифтом** выделены анизотропные минералы, представленные в поворотных сериях, а также наилучшее значение IoU для каждой строки (минерала).

Mineral name (abbreviation)	Baseline	M1	M3	M6
Background (BG)	0.9036	0.906	0.9111	0.9078
Chalcopyrite (Ccp)	0.8266	0.7847	0.8128	0.8126
Galena (Gl)	0.6832	0.669	0.6588	0.6778
Magnetite (Mag)	0.2839	0.3675	0.3226	0.331
Bornite (Br)	0.8055	0.7898	0.7816	0.7932
Pyrite (Py)	0.7929	0.8044	0.8086	0.8088
Sphalerite (Sph)	0.7552	0.7677	0.7822	0.7707
Tennantite (Tnt)	0.6428	0.6992	0.6964	0.704
Pentlandite (Pn)	0.5106	0.4963	0.5092	0.5299
Covellite (Cv)	0.7386	0.722	0.702	0.7247
Hematite (Hem)	0.0774	0.0809	0.0605	0.0832
Arsenopyrite (Apy)	0.7869	0.7778	0.778	0.7817
Marcasite (Mrc)	0.4798	0.3573	0.3626	0.3636
Pyrrhotite (Po)	0.807	0.8143	0.8218	0.7965
Mean IoU \pm std:	0.6496 \pm 0.013	0.6455 \pm 0.0072	0.6434 \pm 0.0089	0.649 \pm 0.0152

- M3: модель, использующая основное изображение и три дополнительных изображения в кросс-поляризованном свете, снятых под углами 0° , 60° и 120° .
- M6: модель с шестью дополнительными входными изображениями, соответствующими снимкам в кросс-поляризованном свете при 0° , 30° , 60° , 90° , 120° и 150° .

Далее проводится сравнение этих конфигураций на исходных данных и на данных после применения предобработки – выравнивания освещенности.

Для оценки устойчивости результатов каждая конфигурация модели обучалась независимо 5 раз с разной инициализацией генератора псевдослучайных чисел. В Таблицах 1 и 2 представлены усреднённые по этим запускам значения метрики Intersection over Union (IoU) для каждого класса, а также среднее значение IoU и его стандартное отклонение (std) по всем классам для каждого эксперимента. Стандартное отклонение характеризует неустойчивость результатов между запусками. На Рис. 8 визуализированы средние значения IoU с интервалом в стандартное отклонение для всех четырех моделей на исходных и скорректированных данных, что позволяет наглядно сравнить как точность, так и стабильность каждой конфигурации.

В таблицах 1 и 2 жирным шрифтом и отдельной линией выделены четыре анизотропных ("мигающих") минерала, присутствующих в поворотных сериях: марказит (Mrc), гематит (Hem), пирротин (Po) и арсенопирит (Apy). Анизотропный ковеллин (Cv) в используемом наборе отсутствует на изображениях в кросс-поляризованном свете, поэтому в эту группу не включен.

Исходя из полученных результатов можно сделать вывод, что использование дополнительных поворотных изображений не приводит к статистически значимому и стабильному улучшению средней точности сегментации (Mean IoU) по сравнению с базовой моделью (Baseline). Напротив, в большинстве конфигураций (M3, M6) наблюдается её снижение.

При этом выравнивание освещенности является важным этапом, который:

- Стабилизирует результаты обучения: стандартные отклонения средних значений метрики IoU (std) для всех моделей существенно уменьшаются (см. Рис. 8) (например, с 0.0274 до 0.013 у Baseline).
- Улучшает или сохраняет среднее качество сегментации: лучший результат на скорректированных данных достигается Baseline (0.6496), превосходя результаты моделей с дополнительными данными и демонстрируя сопоставимую или более высокую точность на большинстве минералов.

Высокие значения стандартного отклонения (см. Рис. 8) (особенно для моделей M1 и M6 на исходных данных) указывают на значительную неустойчивость обучения при использовании дополнительных каналов без предобработки. Это делает такие модели менее надежными и воспроизводимыми на практике.

Несмотря на отсутствие глобального улучшения, для некоторых минералов (например, Magnetite, Hematite) добавление одного дополнительного канала (M1) дает заметный прирост IoU. Это позволяет предположить, что польза от поворотных данных сильно зависит от конкретного минерала и его оптических свойств.

В рамках проведенного эксперимента наиболее эффективной и стабильной является базовая модель ResUNet (Baseline), примененная к данным с выровненной освещенностью. Она демонстрирует наивысшую среднюю точность при минимальном разбросе результатов, что делает её оптимальным выбором для практического применения.

5 Заключение

В работе проведено расширенное исследование влияния использования поворотных изображений и предобработки на сегментацию аншлифов с

анизотропными минералами. Ключевые выводы, полученные на расширенном наборе данных LumenStone с использованием усовершенствованного пайплайна, следующие:

- Использование дополнительных изображений в кросс-поляризованном свете не даёт значимого улучшения средней точности сегментации (Mean IoU) по сравнению с базовой моделью, работающей только с одним изображением в плоско-поляризованном свете.
- Выравнивание освещённости является важным этапом, который существенно повышает устойчивость результатов обучения, снижая вариативность метрик, и способствует росту их средних значений (на 2.19%, 0.17%, 4.06%, 6.1% для моделей Baseline, M1, M3, M6 соответственно).
- Наиболее эффективной и стабильной конфигурацией является базовая модель ResUNet, применяемая к данным после коррекции освещённости.

Таким образом, показано, что для надёжной автоматической сегментации минералов на изображениях шлифов достаточно использовать одно изображение с предварительной коррекцией освещения, что упрощает пайплайн и повышает воспроизводимость метода.

Авторские декларации

Доступность данных и программного кода

Работа выполнялась с использованием суперкомпьютера «МГУ-270» МГУ имени М.В. Ломоносова.

Исходный код доступен по ссылке <https://github.com/KourtKardash/petroscope-polarized>

Набор данных LumenStone доступен по ссылке <https://imaging.cs.msu.ru/en/research/geology/lumenstone>

Конфликт интересов

Отсутствует.

Вклад авторов

С. М. Краснова – написание кода, разработка метода, численные эксперименты, написание текста статьи;

А. В. Хвостиков – написание кода, подготовка набора данных;

Д. М. Коршунов – подготовка набора данных;

А. С. Крылов – разработка метода, написание текста статьи;

Д. В. Сорокин – разработка метода, подготовка плана экспериментов, написание текста статьи.

Список литературы

- [1] From visual diagnostics to deep learning: automatic mineral identification in polished section images / DM Korshunov, AV Khvostikov, GV Nikolaev et al. // *Mining Science and Technology (Russia)*. — 2025. — Vol. 10, no. 3. — Pp. 232–244.
- [2] *Marschallinger Robert*. Automatic mineral classification in the macroscopic scale // *Computers & Geosciences*. — 1997. — Vol. 23, no. 1. — Pp. 119–126.
- [3] *Povarennykh Aleksandr Sergeevich*. Crystal chemical classification of minerals. — Springer, 2014.
- [4] A Novel Method for Evaluation of Ore Minerals Based on Optical Microscopy and Image Analysis: Preliminary Results / Licia Santoro, Marco Lezzerini, Andrea Aquino et al. // *Minerals*. — 2022. — Vol. 12, no. 11. — P. 1348.
- [5] Direct mineralogical imaging of economic ore and rock samples with multi-modal nonlinear optical microscopy / Mung-Chung Kao, Adrian F Pegoraro, David M Kingston et al. // *Scientific reports*. — 2018. — Vol. 8, no. 1. — P. 16917.
- [6] Weakly supervised semantic segmentation for skin cancer via CNN superpixel region response / Yanfei Hong, Guisheng Zhang, Benzhenq Wei et al. // *Multimedia Tools and Applications*. — 2023. — Vol. 82, no. 5. — Pp. 6829–6847.
- [7] *Wilms Christian, Frintrop Simone*. Superpixel-based refinement for object proposal generation // 2020 25th International Conference on Pattern Recognition (ICPR) / IEEE. — 2021. — Pp. 4965–4972.
- [8] Superpixel segmentations for thin sections: Evaluation of methods to enable the generation of machine learning training data sets / Jiaxin Yu, Florian Wellmann, Simon Virgo et al. // *Computers & Geosciences*. — 2023. — Vol. 170. — P. 105232.
- [9] SLIC superpixels compared to state-of-the-art superpixel methods / Radhakrishna Achanta, Appu Shaji, Kevin Smith et al. // *IEEE transactions on pattern analysis and machine intelligence*. — 2012. — Vol. 34, no. 11. — Pp. 2274–2282.
- [10] *Neubert Peer, Protzel Peter*. Compact watershed and preemptive slic: On improving trade-offs of superpixel segmentation algorithms // 2014

- 22nd international conference on pattern recognition / IEEE. — 2014. — Pp. 996–1001.
- [11] *Hu Zhongwen, Zou Qin, Li Qingquan.* Watershed superpixel // 2015 IEEE International Conference on Image Processing (ICIP) / IEEE. — 2015. — Pp. 349–353.
- [12] Automatic identification of minerals in images of polished sections / AV Khvostikov, DM Korshunov, AS Krylov, MA Boguslavskiy // *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.* — 2021. — Vol. 44. — Pp. 113–118.
- [13] *Siddique N Paheding S et al.* U-Net and its variants for medical image segmentation: A review of theory and applications // *IEEE Access.* — 2021. — Vol. 9. — Pp. 82031–82057.
- [14] Adaptive histogram equalization and its variations / Stephen M Pizer, E Philip Amburn, John D Austin et al. // *Computer vision, graphics, and image processing.* — 1987. — Vol. 39, no. 3. — Pp. 355–368.
- [15] *Gangolli Sadjyot Hemant, Fonseca Arnold Johnson Luke, Sonkusare Reena.* Image enhancement using various histogram equalization techniques // 2019 Global Conference for Advancement in Technology (GCAT) / IEEE. — 2019. — Pp. 1–5.
- [16] *Jobson Daniel J, Rahman Zia-ur, Woodell Glenn A.* Properties and performance of a center/surround retinex // *IEEE transactions on image processing.* — 1997. — Vol. 6, no. 3. — Pp. 451–462.
- [17] *Xia Haiying, Liu Min.* Non-uniform illumination image enhancement based on retinex and gamma correction // *Journal of Physics: Conference Series* / IOP Publishing. — Vol. 1213. — 2019. — P. 052072.
- [18] Registration and segmentation of PPL and XPL images of geological polished sections containing anisotropic minerals / DI Razzhivina, DM Korshunov, MA Boguslavsky et al. // *Computational Mathematics and Modeling.* — 2023. — Vol. 34, no. 1. — Pp. 16–26.
- [19] *Lowe David G.* Object recognition from local scale-invariant features // *Proceedings of the seventh IEEE international conference on computer vision* / Ieee. — Vol. 2. — 1999. — Pp. 1150–1157.
- [20] LoFTR: Detector-free local feature matching with transformers / Jiaming Sun, Zehong Shen, Yuang Wang et al. // *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition.* —

2021. — Pp. 8922–8931.

- [21] Data balancing method for training segmentation neural networks / A Kochkarev, A Khvostikov, A Krylov et al. // CEUR Workshop Proceedings: Aachen, Germany. — 2020.
- [22] ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data / Foivos I Diakogiannis, François Waldner, Peter Caccetta, Chen Wu // *ISPRS Journal of Photogrammetry and Remote Sensing*. — 2020. — Vol. 162. — Pp. 94–114.
- [23] Deep residual learning for image recognition / Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun // Proceedings of the IEEE conference on computer vision and pattern recognition. — 2016. — Pp. 770–778.
- [24] *Ronneberger Olaf, Fischer Philipp, Brox Thomas*. U-net: Convolutional networks for biomedical image segmentation // International Conference on Medical image computing and computer-assisted intervention / Springer. — 2015. — Pp. 234–241.
- [25] *Adam Kingma DP Ba J et al*. A method for stochastic optimization // *arXiv preprint arXiv:1412.6980*. — 2014. — Vol. 1412, no. 6.