



Применение машинного обучения к планиметрическому анализу минералов как задаче классификации в различных постановках

Максим Геннадьевич **Шишаев**^{1,2}, Владимир Витальевич **Диковицкий**²

^{1,2} Институт информатики и математического моделирования им. В. А. Путилова КНЦ РАН., Апатиты, Россия

¹ m.shishaev@ksc.ru

Аннотация. Рассматривается проблема автоматизации минералогического анализа планиметрическим методом на примере апатитовых руд различных типов, характерных для Хибинского месторождения.

Цель: изучение эффективности машинного обучения как средства формирования набора признаков и решения задачи классификации в различных постановках при планиметрическом анализе минералов.

Результаты: Выявлены особенности планиметрического анализа как задачи классификации, в частности, определены свойства смежности и однородности классов в пространстве идентифицирующих признаков. Для различных вариантов формальной постановки задач классификации проанализированы потенциальные систематические ошибки определения содержания полезного компонента планиметрическим методом. Экспериментально подтверждена возможность непосредственного использования предобученной сверточной сети ResNet18, без изменений архитектуры и дообучения, для формирования признакового вектора объектов классификации, обеспечивающего хорошую разделимость классов. На примере рассматриваемых руд экспериментально подтверждена высокая эффективность (более 98% точности) применения нейросетевого классификатора и векторизатора ResNet18 для идентификации элементов изображения, относящихся к чистым классам «апатит» / «не-апатит». Высокая точность классификации сохраняется при уменьшении размера ячейки планиметрической сетки вплоть до 2×2 пикселей (78%), а при размере ячейки 20×20 пикселей приближается к 100%.

Исследована эффективность применения нейросетевого подхода к задаче определения удельного содержания полезного компонента в руде. Эксперименты не подтвердили эффективность реализации планиметрического анализа как задачи мягкой классификации без существенных модификаций архитектуры нейросетевого классификатора, однако показали высокую эффективность подхода при мультиклассовой постановке задачи. Абсолютная ошибка в точности определения содержания полезного компонента в последнем случае зависит от количества классов и типа руды и в худшем случае не превышает 6%, что выше точности экспертных оценок опытными рудничными геологами.

Практическая значимость: подход применим для создания недорогих, быстродействующих и эффективных экспресс-анализаторов руд, не требующих специализированного оборудования.

Ключевые слова и фразы: планиметрический минералогический анализ, машинное обучение, ResNet18, классификация

Благодарности: Исследование выполнено в рамках государственного задания ИИММ КНЦ РАН Министерства науки и высшего образования РФ, темы НИР: «Методы и технологии создания интеллектуальных информационных систем для поддержки развития сложных динамических систем с региональной спецификой в условиях неопределённости и риска» (шифр темы FMEZ-2025-0053).

Для цитирования: Шишаев М. Г., Диковицкий В. В. *Применение машинного обучения к планиметрическому анализу минералов как задаче классификации в различных постановках* // Программные системы: теория и приложения. 2025. Т. 16. № 4(67). С. 241–266. https://psta.psisras.ru/read/psta2025_4_241-266.pdf

Введение

Минералогический анализ по визуальным данным представляет собой недорогую и менее затратную по времени альтернативу сложным и дорогостоящим химическим или радиологическим методам анализа. Для определения удельного содержания некоторого минерала по визуальному изображению образца руды используется так называемый площадной анализ, суть которого заключается в оценке относительной площади минерального пятна, соответствующего интересующему нас компоненту (минералу), на плоском изображении образца и экстраполяции полученной площадной оценки на весь его объем. В свою очередь, одним из методов получения площадной оценки является планиметрический метод, в рамках которого на изображение накладывается прямоугольная, обычно квадратная, сетка и затем осуществляется отнесение каждой ячейки сетки к тому или иному компоненту руды.

Таким образом, планиметрический метод минералогического анализа можно рассматривать как задачу классификации, где в роли объектов выступают ячейки планиметрической сетки, а в роли классов – идентифицируемые минералы, присутствующие в образце руды. В «классическом» планиметрическом анализе идентификация объекта с классом осуществляется вручную экспертом-минералогом. Основная идея данной работы заключается в исследовании возможностей автоматизации этого процесса с помощью моделей машинного обучения (МО), которые используются в двух аспектах – для получения абстрактного векторного представления классифицируемого образца и для отнесения последнего к одному из заданных классов.

Мы рассматриваем три варианта постановки задачи автоматизированной идентификации ячеек планиметрической сетки с тем или иным классом: мягкая, бинарная и мультиклассовая классификация. Для каждого варианта предложены формальные постановки и проведены эксперименты по тренировке соответствующих нейросетевых классификаторов и проверке их эффективности при определении удельного содержания минерала в руде. Исследования проводились на примере апатитовых руд Хибинского месторождения различных видов. Задача минералогического анализа рассматривалась в упрощенном виде: определить удельное содержание апатита в образцах руды, представленных цифровыми изображениями. В этом прикладном аспекте задача свелась к бинарной классификации, где положительный класс представляет искомый минерал, а отрицательный класс – все прочие компоненты руды.

Специфической проблемой автоматизации планиметрического анализа является то, что в идеальном случае все ячейки планиметрической сетки должны принадлежать строго одному классу (включать в себя единственный минерал). Однако, на практике, очевидно, это недостижимо — при любом размере сетки наряду с «чистыми» представителями классов, среди ячеек будут присутствовать «пограничные» образцы, относящиеся к более, чем одному классу одновременно. Неизбежна ошибка дискретизации, связанная с тем, что реальные формы минеральных пятен аппроксимируются множеством прямоугольников. Уменьшение размера сетки снижает ошибку дискретизации, попутно снижая возможность учета морфологических и текстурных свойств минералов.

Нужен компромиссный размер сетки, сочетающий точное определение содержания минерала в образце с точной классификацией элементов. Строго говоря, планиметрический анализ является задачей мягкой классификации, когда степень принадлежности объектов классу может лежать в диапазоне от 0 до 1. Крайние значения степени принадлежности соответствуют ячейкам, целиком принадлежащим положительному или отрицательному классам, а промежуточные — ячейкам, лежащим на границе минерального пятна.

Ключом к успешному решению задачи классификации является выбор классифицирующих признаков. Для формирования вектора признаков, обеспечивающего эффективное разделение множества объектов на классы, мы использовали предобученный векторизатор изображений ResNet18, представляющий собой сверточную искусственную нейронную сеть (ИНС), выдающую на выходе 512-компонентный вектор, соответствующий входному изображению [1]. Предварительные эксперименты показали весьма хорошую разделимость рассматриваемых классов объектов в таком признаковом пространстве.

Основной вклад данной работы заключается в анализе различных постановок задач планиметрического анализа как задач классификации и в экспериментальной проверке эффективности их решения методами машинного обучения в применении к площадному анализу изображений, объекты которых обладают свойствами смежности и однородности классов. В частности, в работе адресуются следующие исследовательские вопросы:

- (1) Насколько эффективно использование классификационных признаков, полученных с помощью векторизатора изображений на основе предобученной сверточной ИНС, для идентификации минеральных пятен апатита на изображениях образцов руды в видимом спектре?

- (2) Каковы предельные размеры планиметрической сетки, обеспечивающие эффективное определение удельного содержания апатита по изображению образца руды?
- (3) Какие варианты постановки задачи классификации наиболее эффективны с точки зрения точности определения удельного содержания минерала в руде планиметрическим методом?

Оставшаяся часть статьи организована следующим образом: в первой части кратко рассмотрены задача площадного минералогического анализа и существующий опыт применения машинного обучения для ее решения, а также формальный анализ планиметрического анализа как задачи классификации в различных постановках. Во второй части представлены результаты экспериментов по применению планиметрического анализа в рассматриваемой постановке к определению удельного содержания полезного компонента в апатитовых рудах.

1. Площадной минералогический анализ на базе машинного обучения

Площадной анализ образцов горной породы, впервые предложенный французским минералогом Делессе [2], заключается в измерении площадей минеральных пятен на двумерных изображениях плоских сечений руд или препаратов, изготовленных из разрозненных частиц минералов. При этом, в соответствии с принципом Делессе, предполагается, что площадная оценка эквивалентна объемной. В проекции на задачи машинного зрения, можно, таким образом, выделить два этапа площадного анализа:

- (1) сегментация изображения,
- (2) идентификация сегментов с тем или иным минералом (классификация).

Также возможно совмещение этих двух этапов в рамках семантической сегментации на базе глубоких ИНС [3]. Рассматриваемый в данной работе планиметрический метод площадного анализа, когда изображение разбивается на равные части (ячейки планиметрической сетки) и объектом классификации является отдельно взятая ячейка, можно считать альтернативой методам, основанным на сегментации.

Для автоматизации минералогического анализа в настоящее время широко используется машинное обучение (МО). Вне зависимости от способа выделения идентифицируемых с минералом областей изображения, определяющее влияние на качество последующей классификации

выделенных объектов оказывает выбор набора признаков, которые должны хорошо разделять объекты в соответствующем пространстве. Характерной особенностью ранних и некоторых современных работ в данной области является использование эвристик при формировании признаковых векторов объектов. К таким эвристикам относятся представления изображения в различных цветовых моделях, статистические характеристики изображения, съемка в различных диапазонах и др. Например, авторы [4, 5] используют различные эвристические признаки для выделения сегментов изображения и последующей их идентификации с тем или иным минералом с помощью различных моделей машинного обучения.

Эффективной альтернативой эвристическим подходам к формированию признаков являются предобученные сверточные ИНС, позволяющие выделять абстрактные векторы, наилучшим образом идентифицирующие рассматриваемые классы объектов [6]. Например, в [7] авторы используют сверточную ИНС Yolo для визуальной идентификации каменного угля на фоне пустой породы. В работе [8] с помощью глубокой ИНС авторы решают задачу идентификации образца с минералом из заданного ограниченного набора. В работе [9] авторы исследуют возможность применения различных нейросетевых классификаторов для идентификации коллекционных образцов руды, при этом в качестве векторизатора признаков также используются сверточные ИНС. Отметим, что в упомянутых работах рассматриваются образцы чистых минералов, без примесей, то есть непосредственно решается задача классификации образца, в качестве которого в данном случае выступает изображение целиком.

Одна из первых работ, посвященных минералогическому анализу на основе сегментации изображений, описана в [10]. В ней предложен алгоритм автоматической сегментации цветных изображений горных пород с использованием методов кластерного анализа (K-means) на основе цвето-яркостных признаков. В последующих работах этого исследовательского коллектива [11] к идентифицирующим признакам минералов были добавлены морфологические свойства элементов изображения – площадь и геометрические пропорции минерального пятна. В соответствии с предложенным авторами подходом, на начальном этапе осуществляется кластеризация объектов изображения по цвето-яркостным признакам, а затем с помощью алгоритмов компьютерной графики определяются морфологические признаки полученных кластеров.

В работе [9], наряду с чистыми образцами минералов, рассматриваются случаи, когда минерал представлен в образце породы отдельными

вкраплениями. Применительно к таким случаям в работе предложен оригинальный алгоритм автоматической (zero-shot) сегментации изображения, основанный на GradCAM-методах [12, 13].

Привлекательной стороной методов, основанных на сегментации изображений, является их независимость от способа получения изображения: могут использоваться как обычная фотосъемка в различных световых диапазонах, так и микрофотосъемка или съемка с использованием сканирующих микроскопов. Различные условия съемки позволяют сформировать изображения с максимально контрастным разделением минералов.

Однако, сегментация изображений в применении к задаче площадного минералогического анализа сопряжена с рядом сложностей, к числу которых относятся высокая вариативность текстур и цветов минералов, зашумленность изображений и влияние освещения, сходство минеральных структур и их взаимное перекрытие. Всё это затрудняет создание универсальных алгоритмов сегментации, которые могут работать с разными типами минералов. Работы в направлении решения этой проблемы ведутся на протяжении многих лет, см., например [14, 15], однако кардинального решения пока не найдено.

Значительные успехи в решении задач минералогического анализа на базе машинного обучения связаны с появлением глубоких сверточных нейронных сетей, позволивших решать задачи семантической сегментации. В данном случае осуществляется не только выделение на изображении однородных областей, но и их одновременная идентификация с заданными классами. Однако, применение семантической сегментации на базе глубоких ИНС сопряжено с необходимостью в изображениях высокого качества и в большом объеме тренировочных данных. Это затрудняет получение универсальных недорогих решений, работоспособных на широком спектре различных минералов и не использующих дорогостоящие технические средства.

Например, в работе [16] для выделения минеральных пятен глины на изображениях пород из нефтегазовых месторождений использована свёрточная нейронная сеть U-Net [17]. Это обеспечило довольно высокую точность сегментации (около 92 %), однако при этом использовались изображения, полученные с помощью сканирующего электронного микроскопа (SEM), а эффективность подтверждена только на рассматриваемом минерале. В работе [18] исследована эффективность различных архитектур сверточных сетей в задаче семантической сегментации 2D и 3D изображений мульти-минеральных образцов. Результаты проведенных

авторами экспериментов показали большую вариативность точности сегментации в зависимости от рассматриваемого минерала. В целом, авторы отмечают, что применение глубоких нейронных сетей обеспечивает достаточно высокую точность сегментации, но их обучение требует значительных ресурсов и размеченных данных высокого качества.

Таким образом, решение задачи площадного минералогического анализа по визуальным данным с использованием планиметрического метода с последующей классификацией ячеек планиметрической сетки представляется перспективной менее ресурсоемкой альтернативой подходам на основе сегментации.

В качестве объекта классификации в рассматриваемой задаче выступает ячейка планиметрической сетки (далее в данной статье именуемая также «секцией»). Задача заключается в отнесении каждого элемента множества объектов $X = \{x_1, \dots, x_N\}$ к одному из классов из заданного множества $C = \{c_1, \dots, c_K\}$. В контексте определения содержания полезного компонента¹ в апатитовой руде нас интересует доля минерала «апатит» в общем объеме руды, поэтому, с точки зрения рассматриваемой прикладной задачи, классификация является бинарной – один класс (позитивный) составляют объекты, относящиеся к минеральным пятнам апатита, а второй (негативный) – все остальные объекты. Поскольку, как отмечалось ранее, ячейка сетки может быть пограничной и включать в себя разные минералы, при использовании планиметрического метода анализа мы имеем дело с задачей мягкой классификации, когда для каждого объекта должна быть определена степень принадлежности к классам из множества C

$$c_i(x) \in [0, 1].$$

Одной из гипотез, рассматриваемых в работе, является то, что степень принадлежности к классу в данном случае можно интерпретировать как оценку содержания соответствующего минерала в образце. Тогда удельное содержание k -го минерала будет определяться по формуле

$$(1) \quad p(c_k) = \frac{\sum_i c_k(x_i)}{N}.$$

где N – общее количество объектов классификации (ячеек планиметрической сетки).

¹Ввиду того, что объемное содержание минерала в руде однозначно определяет содержание полезного компонента, в данной статье для упрощения изложения мы будем использовать термин «полезный компонент» также и для обозначения минерала, объемное содержание которого нас интересует.

Рассматриваемую задачу определения полезного компонента в руде (задачу бинарной классификации) можно также интерпретировать как многоклассовую, если считать объекты со степенью принадлежности положительному классу, лежащей в некотором диапазоне $[c_i^{\min}, c_i^{\max}]$, самостоятельным i -м классом. При этом интервал степени принадлежности исходного положительного класса разбивается на K непересекающихся подинтервалов

$$(2) \quad \bigcup_{i=1}^K [c_i^{\min}, c_i^{\max}] = [0, 1], \quad \bigcap_{i=1}^K [c_i^{\min}, c_i^{\max}] = \emptyset.$$

Иными словами, степень принадлежности рассматривается в данном случае как идентифицирующий признак класса. В этом случае удельное содержание искомого минерала (положительного класса) будет определяться по формуле

$$(3) \quad p(c^+) = \frac{\sum_i^K N_i \tilde{c}_i}{N},$$

где N_i – количество идентифицированных объектов (ячеек) i -го класса; $\tilde{c}_i$ – содержание искомого минерала в ячейках i -го класса.

При такой интерпретации наряду с ошибкой классификации на точность влияет и ошибка дискретизации, возникающая из-за того, что объектам, принадлежащим промежуточным классам (со степенью принадлежности к апатиту $c_i(x) \in]0, 1[$), будет приписываться некоторое усредненное содержание искомого минерала, в общем случае отличное от фактического. В предельном случае, когда все объекты класса расположены (по степени принадлежности к положительному классу) на одной из границ интервала $[c_i^{\min}, c_i^{\max}]$, а в качестве $\tilde{c}_i$ выбрана противоположная граница, эта ошибка будет равна $c_i^{\max} - c_i^{\min}$.

Особенностью рассматриваемой здесь задачи является то, что классы в нашем случае не являются произвольными независимыми множествами объектов, но упорядочены друг относительно друга по степени принадлежности их элементов исходному положительному классу. С учетом соотношения (2), назовем это свойством смежности классов в пространстве идентифицирующих признаков. В таком случае между любой парой классов можно определить дистанцию в пространстве c , при этом

$$\tilde{c}_i - \tilde{c}_j \xrightarrow{i \rightarrow K, j \rightarrow 1} 1 = \max.$$

Эта особенность ограничивает возможность увеличения количества классов с целью уменьшения ошибки дискретизации, поскольку это ведет

к возрастанию максимальной дистанции между классами в пространстве s , что негативно сказывается на точности определения содержания искомого минерала.

Обозначим через $\xi = 1 - acc$, где acc – доля правильных ответов (ассигасу), общую ошибку мультиклассового классификатора, а через ξ_{ij} – ошибку классификации объектов j -го класса в пользу i -го класса (доля объектов j -го класса, ошибочно классифицированных как объекты i -го класса). Если при тренировке классификатора не учитывается взаимное расположение классов в пространстве s , то ошибка классификации будет распределена равномерно по классам, то есть

$$\xi_{ij} \approx \frac{\xi}{K-1} \quad \forall i \neq j,$$

где K – общее количество классов.

Тогда ошибка определения содержания полезного компонента (положительного класса) будет равна

$$\delta = \frac{\sum_{i=1}^K \sum_{l \neq i} \frac{\xi}{K-1} \hat{N}_i |\tilde{c}_l - \tilde{c}_i|}{N},$$

где $\hat{N}_i$ – количество объектов, классифицированных как принадлежащих i -му классу, N – общее количество объектов.

Как видим, ошибка зависит от дистанции между классами в пространстве s , что обуславливает дополнительное требование к используемому классификатору на базе машинного обучения: при тренировке классификатора целесообразно использовать штрафную функцию, зависящую от дистанции между классами. Это позволит снизить вероятность ошибок классификатора в пользу классов, удаленных от заданного в пространстве s , и снизить, тем самым, ошибку определения содержания полезного компонента.

Ввиду упомянутой выше особенности задачи, оценку качества классификации в нашем случае целесообразно производить с использованием взвешенных усредненных оценок точности. Обозначим вес ошибки классификации объекта j -го класса в пользу объекта i -го класса как λ_{ij} . С учетом того, что максимально возможная абсолютная ошибка определения удельного содержания полезного компонента (ПК) при ошибочном отнесении объекта к соседнему классу равна $\frac{2}{K}$, вес ошибки будет определяться по формуле

$$\lambda_{ij} = \frac{1 + |i - j|}{K}.$$

Тогда взвешенная оценка количества ложно-положительных срабатываний для i -го класса будет равна

$$FP_i^W = \sum_{j=1}^K \lambda_{ij} FP_{ij},$$

где FP_{ij} – количество объектов j -го класса, ошибочно идентифицированных как объекты i -го класса.

В свою очередь, взвешенная усредненная оценка точности многоклассовой классификации (micro-averaging) будет определяться по формуле

$$Prec^W = \frac{\sum_{i=1}^K TP_i}{\sum_{i=1}^K TP_i + \sum_{i=1}^K FP_i^W}.$$

В контексте задачи классификации изображение апатита обладает однородностью по используемому классифицирующему признаку. Иначе говоря, фрагмент участка изображения, идентифицированного по используемому классифицирующему признаку с классом c_k , будет принадлежать классу c_k с той же степенью, что и исходный участок. Обозначим через $x_i^{d_0}$ объект классификации, полученный путем наложения на исходное изображение планиметрической сетки размером d_0 , а через $x_{ij}^{d_1}$ – разбиение исходного объекта путем наложения более мелкой сетки размером $d_1 < d_0$, то есть

$$\bigcup_j x_{ij}^{d_1} = x_i^{d_0}.$$

Тогда свойство однородности объектов классификации может быть сформулировано следующим образом:

$$(4) \quad \forall x_i^{d_0} : c_k(x_i^{d_0}) \approx 1 \rightarrow c_k(x_{ij}^{d_1}) \approx c_k(x_i^{d_0}),$$

где $c_k(x)$ – степень принадлежности объекта k -му классу.

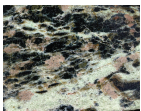
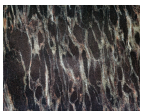
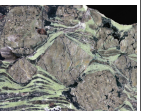
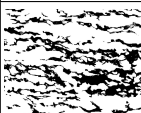
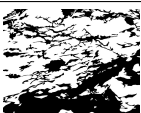
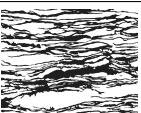
Можно предположить, что существует некоторая нижняя граница размера секции, при которой условие (4) перестает выполняться и используемый для классификации набор признаков перестает давать хороший результат. В предельном случае размер секции может быть равен одному пикселю и тогда признаковый вектор вырождается до используемой цифровой модели изображения, по которой идентифицировать минерал по меньшей мере сложнее, если вообще возможно.

В отношении апатита этот факт подтверждается проведенными экспериментами. Таким образом, значение d , характеризующее размер сетки, является параметром алгоритма классификации.

2. Экспериментальная оценка эффективности планиметрического анализа апатитовых руд с использованием машинного обучения

Для экспериментов использовались пять изображений апатитовой руды Хибинского месторождения различных видов с разной концентрацией полезного компонента и соответствующие им изображения-маски, на которых области, соответствующие апатиту, закрашены в черный цвет, а фон – в белый (таблица 1). Разрешение исходных фото-изображений составляет 300 dpi, размер – 915×709 пикселей, изображения получены обычной фотокамерой. Изображения-маски формировались путем ручной разметки экспертами-минералогами.

ТАБЛИЦА 1. Образцы изображений апатитовой руды, использованные в экспериментах

Вид руды	Апатитовый уртит	Пятнисто-полосчатая	Пятнистая	Линзовидно-полосчатая	Блоковая
Индекс образца	a	b	c	d	e
Фото					
Маска					
Содержание ПК	25%	41%	50%	40%	25%

Для получения набора объектов, используемых в дальнейших экспериментах, на изображения образцов руды и соответствующие им маски накладывалась планиметрическая сетка различных размеров, после чего для каждого объекта рассчитывался показатель фактического содержания полезного компонента как доля черных пикселей в соответствующем изображении-маске в общем числе пикселей изображения. С целью увеличения размера датасетов, в дальнейших экспериментах применялись

приемы аугментации – многократное наложение планиметрической сетки со смещением, а также поворот изображений.

Эталонным набором признаков, однозначно идентифицирующих минерал, по определению, является его химический состав. Очевидно, что химический состав не может быть напрямую определен по изображению образца, поэтому для визуального минералогического анализа необходимы иные наборы признаков, косвенно идентифицирующих тот или иной минерал. Если анализ интерпретируется как задача классификации, такие признаки именуются классификационными. Таким образом, нам необходим вектор признаков, значения которого, с одной стороны, позволяют максимально точно идентифицировать минерал, а с другой – могут быть оценены на основании информации, содержащейся в изображении образца руды.

В нашей работе в качестве классифицирующего признака использовались векторы изображений, полученные с помощью предобученной сверточной остаточной (residual) нейронной сети ResNet18 [1]. Отметим, что в качестве векторизатора могут рассматриваться и другие варианты сверточных сетей [9], однако на задаче распознавания апатита уже ResNet18, с относительно небольшими количеством слоев и размером выходного вектора (512 компонентов), обеспечила достаточно высокое качество результата. Общая схема проведенных экспериментов представлена на рисунке 1.

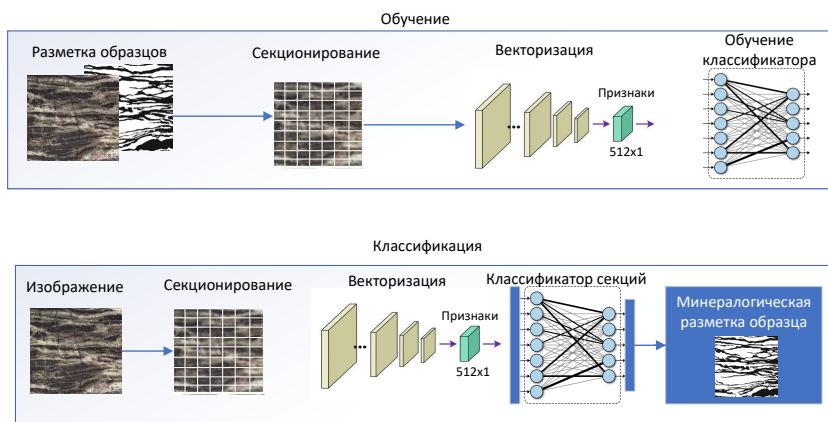


Рисунок 1. Общая схема экспериментов

В первой серии экспериментов изучалась эффективность использования векторизатора ResNet18 для идентификации ячеек планиметрической

сетки с апатитом и прочими компонентами апатитовых руд различных типов. Задача интерпретировалась как задача бинарной классификации — для каждой ячейки планиметрической сетки необходимо определить принадлежность к положительному (apatit) или отрицательному (не-apatit) классам.

Для оценки возможности применения простых методов классификации, основанных на некоторой метрике в пространстве признаков, был проведен эксперимент по оценке близости векторов изображений различных размеров, полученных из фотографии одного размеченного образца. В качестве исходных данных взято изображение апатит-содержащей руды и соответствующая ей маска, представленные на рисунке 2.

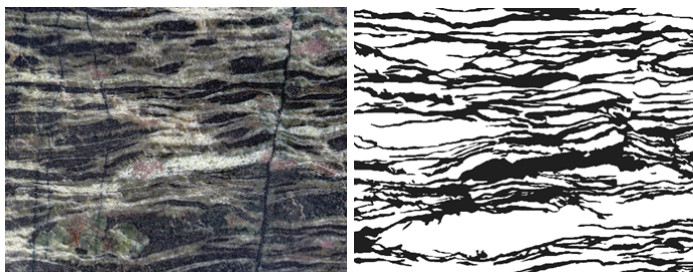


РИСУНОК 2. Использованное в эксперименте изображение апатитовой руды (слева) и соответствующая ему маска (справа)

На изображение была наложена планиметрическая сетка размером 50×50 пикселей со смещением, в результате получен набор из 902 объектов положительного и 4573 объектов отрицательного классов. Ячейки сетки, лежащие на границах минеральных пятен, из рассмотрения исключены. В качестве метрики близости объектов использовалось косинусное расстояние между векторами секций изображения, полученными с помощью ResNet18 (слой «Global Average Pool»).

Далее случайным образом был выбран эталонный объект класса «apatit» и относительно его векторного представления рассчитаны косинусные расстояния до объектов положительного и отрицательного классов. Диапазон расстояний между эталоном и объектами положительного класса составил: 0,75–0,98 со средним значением 0,8855. Аналогичный диапазон для объектов отрицательного класса составил 0,46–0,91 со средним значением 0,7523.

Мы видим, что объекты отрицательного и положительного классов могут быть достаточно близки по косинусному расстоянию. Таким образом,

непосредственное использование в качестве классификационного признака дистанции в пространстве векторов, полученных с помощью ResNet18, не эффективно. Вместе с тем, хорошая разделимость объектов, полученная с помощью РСА² (в силу ограниченности объема, эти эксперименты в данной статье не рассмотрены), указывает на возможность тренировки достаточно точного нейросетевого классификатора.

Для проверки этой возможности проведен эксперимент по бинарной классификации планиметрических ячеек с помощью ИНС прямого распространения. Использовались несколько эталонных (размеченных) образцов апатитовой руды с различной морфологией, классификатор тренировался на изображениях одних типов руд и тестировался на других. Это обеспечило наряду с кроссвалидацией проверку на внешних данных.

Задача распознавания апатита на изображениях руды также решалась как задача бинарной классификации фрагментов изображения. Для получения образцов данных, используемых для обучения модели, на исходные изображения накладывалась планиметрическая сетка с ячейкой 10×10 пикселей. Каждый образец затем получал бинарную метку («апатит/не-апатит») в соответствии с имеющейся маской. Пограничные образцы, маска которых содержала количество черных или белых пикселей менее заданного порога, из выборки исключались (рисунок 3).

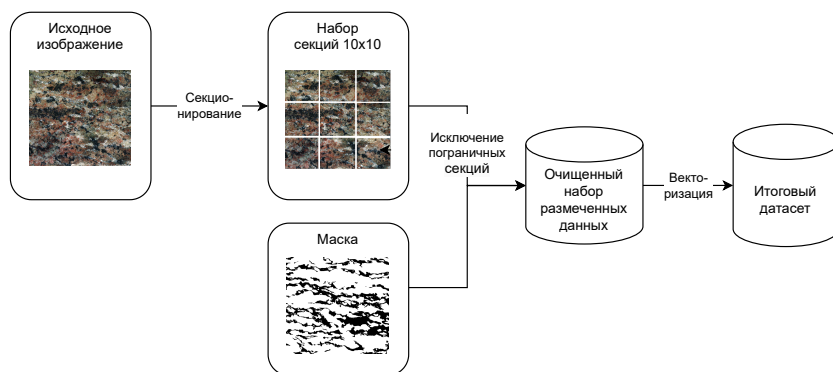


Рисунок 3. Схема формирования датасета для обучения бинарного классификатора

В соответствии с содержанием ПК (см. таблицу 1), для большинства исходных изображений имело место разбалансирование выборки

²Метод главных компонент (Principal Component Analysis, PCA)

(преобладание класса «не-апатит» в образцах). Для балансировки выборки в проведенных экспериментах использовалась субдискретизация (under-sampling) [19] методом случайного сокращения мажоритарного класса.

Классификатор тренировался на трех разных выборках, полученных из образцов (а), (b) и на всех изображениях, за исключением образца (с). Использовалась ИНС прямого распространения с пятью полносвязными слоями: начальный слой принимает 512-мерные векторы ResNet18, с последующими нелинейными преобразованиями через слои с активациями softmax, relu и sigmoid. Выходной слой имеет одну нейронную единицу с активацией sigmoid, интерпретируемую как вероятность принадлежности объекта положительному классу. Сеть обучена с использованием функции потерь binary-crossentropy и оптимизатора Adam. Общее количество тренируемых параметров: 618660. Кривые точности и функции потерь на обучающей и валидационной выборке представлены на рисунке 4.

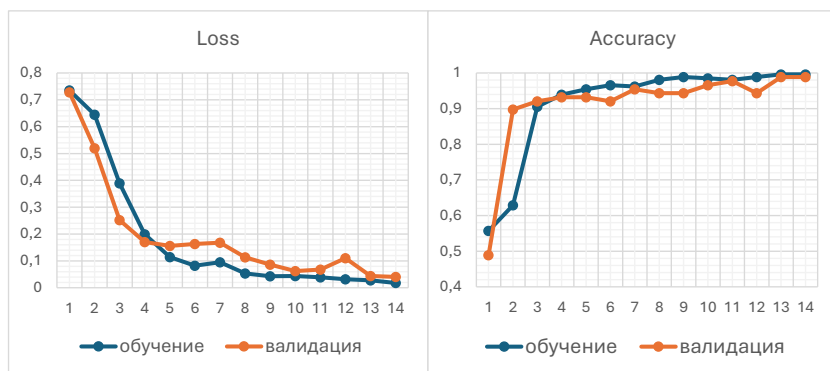
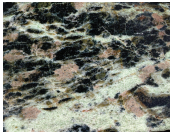
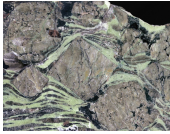
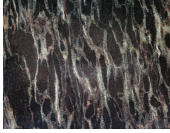


РИСУНОК 4. Кривые точности и функции потерь на обучающей и валидационной выборке

Наряду с валидационной частью выборки, полученные классификаторы тестировались также на внешних данных, в качестве которых выступали изображения образцов руды, не участвовавшие в обучении. Результаты тестирования (оценивалась точность классификации) представлены в таблице 2.

Результаты показывают, что нейросетевой классификатор хорошо справляется с идентификацией апатита в рамках одного типа руды или некоторого набора руд со схожими морфологическими признаками. Вместе

ТАБЛИЦА 2. Результаты экспериментов с классификатором, обученным на образцах (а), (b) и на всех изображениях

Образец		Точность (Precision), %		
		Классификатор, обученный на образце (b)	Классификатор, обученный на образце (a)	Классификатор, обученный на всех образцах, исключая (c)
b		99,87^a	93,776	98,5
d		95,071	85,417	
e		93,532	51,087	
a		59,77	98,64	
c		38,552	36,471	98,824

^a Жирным выделены оценки точности на валидационной части обучающей выборки.

с тем, добавление в обучающую выборку изображений различных типов руд не снижает точность классификации.

Для оценки влияния на эффективность классификации размера ячейки планиметрической сетки была проведена тренировка и тестирование классификаторов на датасетах, сформированных аналогичным описанному выше образом из всех имеющихся размеченных изображений, с размерами ячейки 2×2, 5×5, 10×10 и 20×20 пикселей. Выборки с малым размером были искусственно расширены путем смещения планиметрической сетки и вращения ячеек. Для исключения влияния размера датасетов для

различных размеров ячейки используемые для тренировки наборы данных были выровнены путем случайного отбора равного количества объектов положительного и отрицательного классов (по 50 тысяч экземпляров каждого класса). Результаты экспериментов приведены в таблице 3.

Таблица 3. Эффективность бинарной классификации при различных размерах ячейки

Размер ячейки	2×2	5×5	10×10	20×20
Точность классификации на валидационной части выборки	0,7771	0,9023	0,9863	0,9994

Как видим, точность классификации предсказуемо растет с увеличением размера ячейки. Это можно объяснить более полной представленностью морфологических признаков минерала при большем размере ячейки (читай – большем размере изображения минерала). Верхнее ограничение размера ячейки, на практике, обуславливается физическим размером минеральных пятен.

Таким образом, использование нейросетевого векторизатора ResNet18 (без дообучения и модификации архитектуры) обеспечивает возможность эффективной бинарной классификации минерала апатит. Точность идентификации апатита приближается к единице уже при размере ячейки 20×20 пикселей и является достаточно высокой и при меньших размерах ячейки.

Для определения удельного содержания ПК в образце руды в целом необходимо учитывать как «чистые» представители позитивного и негативного классов, так и «пограничные» объекты. Далее в статье рассмотрены два варианта решения этой задачи: (1) интерпретация значений выходов бинарного классификатора как степеней принадлежности к классу и (2) мультиклассовая классификация.

Для проверки возможности непосредственного использования бинарного классификатора в задаче мягкой классификации с последующей интерпретацией значения принадлежности классу как удельного содержания искомого минерала использовалось размеченное изображение апатита повышенного разрешения с размером секции 20×20 пикселей. Бинарному нейросетевому классификатору, натренированному на размеченном датасете, включающем объекты положительного и отрицательного классов, были предъявлены 99 пограничных объектов с различным содержанием минерала, представляющих промежуточный класс. В результате, классификатор показал тенденцию к однозначному отнесению объекта

к положительному классу при удельном содержании целевого минерала в нем 50% и более. Расчетное значение предсказанного содержания ПК в рассмотренных пограничных объектах, вычисленное по формуле (1), составило 0,84, при истинном значении 0,55. Таким образом, возможность непосредственной интерпретации выходных значений бинарного классификатора как степеней принадлежности классу (при использованной архитектуре и технологии обучения нейросетевого классификатора) не подтвердилась.

Для исследования эффективности решения задачи в мультиклассовой постановке осуществлялась тренировка многоклассовых нейросетевых классификаторов на имеющихся размеченных изображениях апатитовой руды, представленных в таблице 1. Множество объектов классификации формировалось путем наложения на изображения квадратной планиметрической сетки размером 10×10 пикселей. Были проведены две серии экспериментов с количеством классов 4 и 10. Границы классов определялись путем равномерного деления интервала удельного содержания ПК на 4 и 10 отрезков, соответственно. В полученных выборках доминировали объекты граничных классов, поэтому была произведена балансировка обучающей выборки методом чересстрочного удаления данных избыточного класса. Для увеличения количества объектов в миноритарных классах обучающей выборки также использовалась аугментация данных путем смещения планиметрической сетки на 5 пикселей, за счет чего обеспечено удвоение размера классов с малым количеством объектов. Общее число объектов обучающей выборки составило 12892.

В таблице 4 представлены параметры качества классификации, полученные в экспериментах с четырьмя классами. Можно отметить

Таблица 4. Параметры качества классификации для четырёх классов

Класс	1	2	3	4	Macro Average	Micro Average
Содержание ПК	0,75–1	0,5–0,75	0,25–0,5	0–0,25		
Recall	0,635	0,718	0,678	0,738	0,692	0,691
Precision	0,813	0,340	0,335	0,707	0,549	0,453
Accuracy	0,872	0,531	0,575	0,877	0,691	n/a
Weighted Precision	0,842	0,629	0,751	0,896	0,780	0,755

ожидаемо худшую точность классификации промежуточных классов (класс 2 и класс 3) в сравнении с пограничными классами. В целом низкая усредненная точность классификации, вычисленная без учета

специфики классов в рассматриваемой задаче, и достаточно высокая взвешенная оценка точности объясняется тем, что большая часть ошибок классификатора происходит между соседними классами.

Для экспериментов с 10 классами получены схожие параметры качества классификации.

Для проверки эффективности многоклассовой классификации в задаче определения удельного содержания ПК в руде обученным классификаторам были предъявлены полные выборки объектов, полученные путем наложения планиметрической сетки на имеющиеся образцы изображений. Размер полных выборок составил 31850 и 126700 объектов для экспериментов с четырьмя и десятью классами, соответственно (большой размер выборки для 10 классов обеспечен аугментацией набора данных путем последовательного смещения планиметрической сетки на 5 пикселей).

Предсказанный удельный объем ПК по всему набору объектов вычислялся по формуле (3). Значения усредненного удельного содержания ПК для каждого класса (параметр $\tilde{c}_i$) приведены в таблице 5.

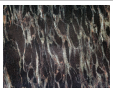
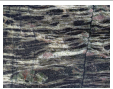
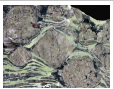
Таблица 5. Параметр $\tilde{c}_i$ для различных классов

Класс	1	2	3	4	5	6	7	8	9	10
4 класса	0,875	0,625	0,375	0,125	-	-	-	-	-	-
10 классов	0,95	0,85	0,75	0,65	0,55	0,45	0,35	0,25	0,15	0,05

Фактическое содержание ПК по всей выборке составляло 37,24%. Предсказанные значения содержания ПК при использовании 4-классового и 10-классового классификаторов составили 45,6% и 39,54% соответственно. Величина абсолютной ошибки, таким образом, составила 8,5% для 4-классового и 2,3% для 10-классового классификаторов. В таблице 6 представлены результаты применения 10-классового классификатора для определения удельного содержания ПК в каждом образце руды и по всей выборке в целом.

По экспертным оценкам, опытный рудничный геолог способен без использования инструментальных методов анализа оценить удельное содержание апатита в руде с абсолютной ошибкой до 8%. Таким образом, многоклассовая классификация обеспечивает точность определения содержания ПК, сопоставимую с точностью экспертных оценок, а при достаточно большом количестве классов—значительно превышающую таковую.

ТАБЛИЦА 6. Результаты применения многоклассовой классификации для определения удельного содержания ПК в образцах

Индекс	a	b	c	d	e	a-e
Образец						
Фактическое содержание ПК, %	31,45	39,14	50,30	40,63	24,69	37,24
Предсказанное содержание ПК, %	36,44	43,66	48,59	38,45	30,55	39,54
Абсолютная ошибка, %	4,99	4,52	1,71	2,18	5,85	2,30

Заключение

В работе исследована возможность и эффективность реализации планиметрического минералогического анализа как задачи классификации с использованием методов машинного обучения. Предложены и проанализированы, в применении к рассматриваемой прикладной задаче, формальные постановки и потенциальная эффективность задач мягкой, бинарной и мультиклассовой классификации. Экспериментальная проверка эффективности предложенных подходов осуществлена на примере пяти различных видов апатитовых руд Хибинского месторождения.

Эксперименты показали высокую эффективность использования предобученной ИНС ResNet18 для извлечения из изображений, полученных обычной камерой, классификационных признаков целевого минерала. Причем использованный признаковый вектор обеспечивает высокую точность классификации даже при относительно малых размерах ячейки планиметрической сетки и существенно деградирует лишь при размере секции 2×2 пикселя.

Реализация планиметрического анализа как задачи классификации в различных постановках показала, что применение мультиклассового классификатора обеспечивает точность определения удельного содержания целевого минерала в образцах, сопоставимую с точностью экспертных оценок опытным рудничным геологом. При этом планиметрический подход к минералогическому анализу руд представляется более универсальным и менее требовательным к качеству обучающих данных и сложности используемых моделей и технических средств в сравнении с подходами на основе семантической сегментации изображений.

Несмотря на то, что в работе рассмотрен лишь один целевой минерал – апатит, результаты могут быть распространены и на задачи анализа других видов минеральных руд, изображения которых обладают свойствами однородности и смежности классов, сформулированными в данной статье. Следует отметить, что подобными свойствами могут обладать и изображения, анализируемые в контексте других прикладных задач, отличных от минералогического анализа. Это открывает возможности для использования аналогичного (планиметрического) подхода к решению широкого спектра задач анализа изображений, где требуется определять удельную площадь визуальных объектов некоторого класса. При этом могут использоваться любые источники исходных данных – от специализированных микроскопов до данных дистанционного зондирования Земли.

Список использованных источников

- [1] He K., Zhang X., Ren S., Sun J. *Deep residual learning for image recognition* // *2016 IEEE Conference on Computer Vision and Pattern Recognition* (27–30 June 2016, Las Vegas, NV, USA).– IEEE.– 2016.– ISBN 9781467388528.– Pp. 770–778.
↑^{243, 252}
- [2] Weibel E. R., Elias H. *Introduction to stereologic principles* // *Quantitative Methods in Morphology / Quantitative Methoden in der Morphologie*, eds. E. R. Weibel, H. Elias, Berlin–Heidelberg: Springer.– 1967.– ISBN 978-3-642-50130-2.– Pp. 89–98.
↑²⁴⁴
- [3] Mo Y., Wu Y., Yang X., Liu F., Liao Y. *Review the state-of-the-art technologies of semantic segmentation based on deep learning* // *Neurocomputing*.– 2022.– Vol. **493**.– Pp. 626–646. ↑²⁴⁴
- [4] Maitre J., Boucharde K., Bédard L. P. *Mineral grains recognition using computer vision and machine learning* // *Computers & Geosciences*.– 2019.– Vol. **130**.– Pp. 84–93. ↑²⁴⁵
- [5] Чернявский А. В., Степенчиков Д. Г. *Метод сегментации изображения для подсчета процентного содержания минералов с помощью авторской программы* // Труды Ферсмановской научной сессии ГИ КНЦ РАН.– 2021.– Т. **18**.– С. 414–418. ↑²⁴⁵
- [6] Li Z., Liu F., Yang W., Peng S., Zhou J. *A survey of convolutional neural networks: analysis, applications, and prospects* // *IEEE Transactions on Neural Networks and Learning Systems*.– 2022.– Vol. **33**.– No. 12.– Pp. 6999–7019. ↑²⁴⁵
- [7] Wang W., Zhao Y., Xue Z. *YOLOv8-Coal: a coal-rock image recognition method based on improved YOLOv8* // *PeerJ Computer Science*.– 2024.– Vol. **10**.– id. e2313. ↑²⁴⁵
- [8] Ji X., Liang K., Yang Y., Yang M., He M., Zhang Z., Zeng S., Wang Y. *Identifying minerals from image using out-of-distribution artificial intelligence-based model* // *Minerals*.– 2024.– Vol. **14**.– No. 6.– id. 627.– 10 pp. ↑²⁴⁵

- [9] Nesteruk S., Agafonova Ju., Pavlov I., Gerasimov M., Latyshev N., Dimitrov D., Kuznetsov A., Kadurin A., Plechov P. *MineralImage5k: A benchmark for zero-shot raw mineral visual recognition and description* // Computers & Geosciences.– 2023.– Vol. **178**.– id. 105414.  [↑245, 252](#)
- [10] Baklanova O., Shvets O. *Cluster analysis methods for recognition of mineral rocks in the mining industry* // 2014 4th International Conference on Image Processing Theory, Tools and Applications (IPTA) 2014 4th International Conference on Image Processing Theory, Tools and Applications (IPTA) (14–17 October 2014, Paris, France).– IEEE.– 2014.– ISBN 9781479964642.– Pp. 1–5.  [↑245](#)
- [11] Baklanova O. E., Baklanov M. A. *Methods and algorithms of image recognition for mineral rocks in the mining industry* // Advances in Swarm Intelligence.– V. II, ICSI 2016 (June 25–30, 2016, Bali, Indonesia), Lecture Notes in Computer Science.– vol. **9713**, eds. Y. Tan, Y. Shi, L. Li, Cham: Springer.– 2016.– ISBN 978-3-319-41009-8.– Pp. 253–262.  [↑245](#)
- [12] Chattopadhyay A., Sarkar A., Howlader A., Balasubramanian V. N. *Grad-CAM++: generalized gradient-based visual explanations for deep convolutional networks* // 2018 IEEE Winter Conference on Applications of Computer Vision (WACV) (12–15 March 2018, Lake Tahoe, NV, USA).– IEEE.– 2018.– ISBN 9781538648872.– Pp. 839–847.  [↑246](#)
- [13] Srinivas S., Fleuret F. *Full-gradient representation for neural network visualization* // Proceedings of the 33rd International Conference on Neural Information Processing Systems (December 8–14, 2019, Vancouver BC, Canada), Advances in Neural Information Processing Systems.– vol. **32**.– Curran Associates, Inc..– 2019.– ISBN 9781713807933.– Pp. 4124–4133.– id. 371.  [↑246](#)
- [14] Launeau P., Cruden A., Bouchez J. *Mineral recognition in digital images of rocks: a new approach using multichannel classification* // Canadian Mineralogist.– 1994.– Vol. **32**.– No. 4.– Pp. 919–933.  [↑246](#)
- [15] Song W., Dong L., Zhao X., Xia J., Liu T., Shi Y. *Review of nodule mineral image segmentation algorithms for deep-sea mineral resource assessment* // Computers, Materials and Continua.– 2022.– Vol. **73**.– No. 1.– Pp. 1649–1669.  [↑246](#)
- [16] Chen Z., Liu X., Yang J., Little E., Zhou Yu. *Deep learning-based method for SEM image segmentation in mineral characterization, an example from Duvernay Shale samples in Western Canada Sedimentary Basin* // Computers & Geosciences.– 2020.– Vol. **138**.– id. 104450.  [↑246](#)
- [17] Ronneberger O., Fischer O., Brox T. *U-Net: Convolutional networks for biomedical image segmentation* // Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015.– V. III (October 5–9, 2015, Munich, Germany), Lecture Notes in Computer Science.– vol. **9351**, Cham: Springer.– 2015.– ISBN 978-3-319-24574-4.– Pp. 234–241.  [% arXiv:1505.04597](#)  [↑246](#)
- [18] Wang Y. D., Shabaninejad M., Armstrong R. T., Mostaghimi P. *Deep neural networks for improving physical accuracy of 2D and 3D multi-mineral segmentation of rock micro-CT images* // Applied Soft Computing.– 2021.– Vol. **104**.– id. 107185.  [↑246](#)

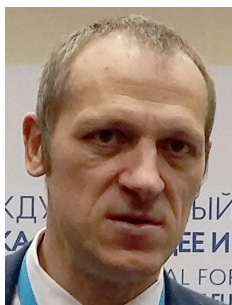
- [19] Yen S.-J., Lee Y.-S. *Under-sampling approaches for improving prediction of the minority class in an imbalanced dataset* // *Intelligent Control and Automation, International Conference on Intelligent Computing, ICIC 2006* (Kunming, China, August 16–19, 2006), *Lecture Notes in Control and Information Sciences.*– vol. **344**, eds. D.-S. Huang, K. Li, G. W. Irwin, Berlin–Heidelberg: Springer.– 2006.– ISBN 978-3-540-37256-1.– Pp. 731–740.  [↑255](#)

Поступила в редакцию 12.10.2025;
одобрена после рецензирования 13.10.2025;
принята к публикации 29.10.2025;
опубликована онлайн 14.11.2025.

Рекомендовал к публикации

к.т.н. Е. П. Куршев

Информация об авторах:



Максим Геннадьевич Шишаев

д.т.н., гл. научн. сотр. Института информатики и математического моделирования имени В. А. Путилова КНЦ РАН. Научные интересы: интеллектуальные информационные системы, прикладные модели искусственного интеллекта, интеллектуальный анализ текстов, инженерия знаний



0000-0001-7070-7878

e-mail: m.shishaev@ksc.ru



Владимир Витальевич Диковицкий

к.т.н., ст. научн. сотр. Института информатики и математического моделирования имени В. А. Путилова КНЦ РАН. Научные интересы: Искусственный интеллект и машинное обучение, обработка информации и анализ данных, вычислительная лингвистика и семантический анализ



0000-0003-0329-9979

e-mail: v.dikovitskiy@ksc.ru

Вклад авторов: М. Г. Шишаев – 50% (идея, методология, формальный анализ, написание черновой версии, доработка и редактирование, администрирование); В. В. Диковицкий – 50% (идея, программное обеспечение, валидация, расследование, сбор материала, визуализация).

Декларация об отсутствии личной заинтересованности: *благополучие авторов не зависит от результатов исследования.*



Machine learning-based mineralogical analysis using planimetric method

Maksim Gennad'evich **Shishaev**¹, Vladimir Vital'evich **Dikovitsky**²

^{1,2}Putilov Institute for Informatics and Mathematical Modeling of the Kola Science Centre of the Russian Academy of Sciences, Apatity, Russia

¹m.shishaev@ksc.ru

Abstract. The problem of automation of mineralogical analysis by the planimetric method is considered using the analysis of various types of apatite ores typical of the Khibiny deposit as an example.

Purpose: To study the efficiency of machine learning as a means for forming a features vector and solving the classification problem in various formulations during planimetric analysis of minerals.

Results: The study revealed the features of planimetric analysis as a classification problem, the properties of adjacency and homogeneity of classes in the space of identifying features are determined. Potential systematic errors in determining the content of the valuable component by the planimetric method are analyzed for various formulations of the classification problems. Experiments confirmed the possibility of directly using the pre-trained convolutional network ResNet-18 for forming a feature vector of classification objects, ensuring good separability of classes. Using the example of the ores under consideration, the high efficiency (over 98% precision) of the neural network classifier and vectorizer ResNet-18 for identifying image elements related to the pure classes "apatite" / "non-apatite" is experimentally confirmed. High classification precision is maintained with planimetric grid cell sizes down to 2×2 pixels (78%), and approaches 100% at 20×20 pixels.

The effectiveness of a neural network approach to determining the specific grade of a valuable component in ore was studied. Experiments did not confirm the effectiveness of implementing planimetric analysis as a soft classification problem without significant modifications to the classifier's architecture. However, they demonstrated the high effectiveness of the approach in a multi-class setting of the problem. The absolute error in determining the grade of a valuable component in the latter case depends on the number of classes and the ore type and, in the worst case, does not exceed 6%, which is higher than the accuracy of expert assessments by experienced mine geologists.

Practical relevance: the approach is applicable to the development of inexpensive, fast, and efficient express ore analyzers that do not require specialized equipment. (*In Russian*).

Key words and phrases: planimetric mineralogical analysis, machine learning, ResNet18, classification

2020 *Mathematics Subject Classification:* 68T45; 00A99

Acknowledgments: The study was carried out within the framework of the Putilov Institute for Informatics and Mathematical Modeling of the Kola Science Centre of the Russian Academy of Sciences state assignment of the Ministry of Science and Higher Education of the Russian Federation, research topic: "Methods and technologies for creating intelligent information systems to support the development of complex dynamic systems with regional specificity in conditions of uncertainty and risk" (topic code FMEZ-2025-0053)

For citation: Maksim G. Shishaev, Vladimir V. Dikovitsky. *Machine learning-based mineralogical analysis using planimetric method*. Program Systems: Theory and Applications, 2025, 16:4(67), pp. 241–266. (*In Russ.*). https://psta.psiras.ru/read/psta2025_4_241-266.pdf

References

- [1] K. He, X. Zhang, S. Ren, J. Sun. “Deep residual learning for image recognition”, *2016 IEEE Conference on Computer Vision and Pattern Recognition* (27–30 June 2016, Las Vegas, NV, USA), IEEE, 2016, ISBN 9781467388528, pp. 770–778. 
- [2] E. R. Weibel, H. Elias. “Introduction to stereologic principles”, *Quantitative Methods in Morphology / Quantitative Methoden in der Morphologie*, eds. E.R. Weibel, H. Elias, Springer, Berlin–Heidelberg, 1967, ISBN 978-3-642-50130-2, pp. 89–98. 
- [3] Y. Mo, Y. Wu, X. Yang, F. Liu, Y. Liao. “Review the state-of-the-art technologies of semantic segmentation based on deep learning”, *Neurocomputing*, **493** (2022), pp. 626–646. 
- [4] J. Maitre, K. Bouchard, L. P. Bédard. “Mineral grains recognition using computer vision and machine learning”, *Computers & Geosciences*, **130** (2019), pp. 84–93. 
- [5] A. V. Chernyavskij, D. G. Stepshshikov. “Image segmentation method for counting the percentage of the minerals by author’s program”, *Trudy Fersmanovskoj nauchnoj sessii GI KNC RAN*, **18** (2021), pp. 414–418 (in Russian). 
- [6] Z. Li, F. Liu, W. Yang, S. Peng, J. Zhou. “A survey of convolutional neural networks: analysis, applications, and prospects”, *IEEE Transactions on Neural Networks and Learning Systems*, **33**:12 (2022), pp. 6999–7019. 
- [7] W. Wang, Y. Zhao, Z. Xue. “YOLOv8-Coal: a coal-rock image recognition method based on improved YOLOv8”, *PeerJ Computer Science*, **10** (2024), id. e2313. 
- [8] X. Ji, K. Liang, Y. Yang, M. Yang, M. He, Z. Zhang, S. Zeng, Y. Wang. “Identifying minerals from image using out-of-distribution artificial intelligence-based model”, *Minerals*, **14**:6 (2024), id. 627, 10 pp. 
- [9] S. Nesteruk, Ju. Agafonova, I. Pavlov, M. Gerasimov, N. Latyshev, D. Dimitrov, A. Kuznetsov, A. Kadurin, P. Plechov. “MineralImage5k: A benchmark for zero-shot raw mineral visual recognition and description”, *Computers & Geosciences*, **178** (2023), id. 105414. 
- [10] O. Baklanova, O. Shvets. “Cluster analysis methods for recognition of mineral rocks in the mining industry”, *2014 4th International Conference on Image Processing Theory, Tools and Applications (IPTA) 2014 4th International Conference on Image Processing Theory, Tools and Applications (IPTA)* (14–17 October 2014, Paris, France), IEEE, 2014, ISBN 9781479964642, pp. 1–5. 
- [11] O. E. Baklanova, M. A. Baklanov. “Methods and algorithms of image recognition for mineral rocks in the mining industry”, *Advances in Swarm Intelligence*. V. II, ICSI 2016 (June 25–30, 2016, Bali, Indonesia), Lecture Notes in Computer Science, vol. **9713**, eds. Y. Tan, Y. Shi, L. Li, Springer, Cham, 2016, ISBN 978-3-319-41009-8, pp. 253–262. 

- [12] A. Chattopadhyay, A. Sarkar, A. Howlader, V.N. Balasubramanian. “Grad-CAM++: generalized gradient-based visual explanations for deep convolutional networks”, *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)* (12–15 March 2018, Lake Tahoe, NV, USA), IEEE, 2018, ISBN 9781538648872, pp. 839–847. 
- [13] S. Srinivas, F. Fleuret. “Full-gradient representation for neural network visualization”, *Proceedings of the 33rd International Conference on Neural Information Processing Systems* (December 8–14, 2019, Vancouver BC, Canada), Advances in Neural Information Processing Systems, vol. **32**, Curran Associates, Inc., 2019, ISBN 9781713807933, pp. 4124–4133, id. 371. 
- [14] P. Launeau, A. Cruden, J. Bouchez. “Mineral recognition in digital images of rocks: a new approach using multichannel classification”, *Canadian Mineralogist*, **32**:4 (1994), pp. 919–933. 
- [15] W. Song, L. Dong, X. Zhao, J. Xia, T. Liu, Y. Shi. “Review of nodule mineral image segmentation algorithms for deep-sea mineral resource assessment”, *Computers, Materials and Continua*, **73**:1 (2022), pp. 1649–1669. 
- [16] Z. Chen, X. Liu, J. Yang, E. Little, Yu. Zhou. “Deep learning-based method for SEM image segmentation in mineral characterization, an example from Duvernay Shale samples in Western Canada Sedimentary Basin”, *Computers & Geosciences*, **138** (2020), id. 104450. 
- [17] O. Ronneberger, O. Fischer, T. Brox. “U-Net: Convolutional networks for biomedical image segmentation”, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. V. III (October 5–9, 2015, Munich, Germany), Lecture Notes in Computer Science, vol. **9351**, Springer, Cham, 2015, ISBN 978-3-319-24574-4, pp. 234–241.   1505.04597
- [18] Y. D. Wang, M. Shabaninejad, R. T. Armstrong, P. Mostaghimi. “Deep neural networks for improving physical accuracy of 2D and 3D multi-mineral segmentation of rock micro-CT images”, *Applied Soft Computing*, **104** (2021), id. 107185. 
- [19] S.-J. Yen, Y.-S. Lee. “Under-sampling approaches for improving prediction of the minority class in an imbalanced dataset”, *Intelligent Control and Automation*, International Conference on Intelligent Computing, ICIC 2006 (Kunming, China, August 16–19, 2006), Lecture Notes in Control and Information Sciences, vol. **344**, eds. D.-S. Huang, K. Li, G. W. Irwin, Springer, Berlin–Heidelberg, 2006, ISBN 978-3-540-37256-1, pp. 731–740. 