

Проблемы искусственного интеллекта. 2026. № 1 (40). С. 202-214
Problems of Artificial Intelligence. 2026;1(40):202-214.

Системный анализ, управление и обработка информации, статистика
Научная статья

УДК 004.932

doi: 10.24412/2413-7383-2026-1-40-202-214

К. И. Решетников¹, М. С. Глухов², М.В. Ронкин¹

¹ФГАОУ ВО «УрФУ имени первого Президента России Б.Н. Ельцина»
620002, г. Екатеринбург, ул. Мира, д. 19

²Институт геологии и геохимии имени акад. А.Н. Заварицкого Уральского отделения
Российской академии наук (ИГГ УрО РАН) 620110, Екатеринбург, ул. Академика Вонсовского;
Институт геологии и нефтегазовых технологий Казанского (Приволжского) федерального
университета (К(П)ФУ), 420008, Россия, РТ, г. Казань, ул. Кремлевская, д.18.

ИССЛЕДОВАНИЕ ПОДХОДОВ КОМПЬЮТЕРНОГО ЗРЕНИЯ ДЛЯ ОБНАРУЖЕНИЯ ОБЪЕКТОВ В ЗАДАЧАХ ФРАГМЕНТАЦИИ ОБЛОМКОВ ГОРНЫХ ПОРОД

K. I. Reshetnikov¹, M. S. Glukhov², M. V. Ronkin¹

¹Ural Federal University named after the First President of Russia B.N. Yeltsin
620002, Yekaterinburg, Mira St., 19

²Zavaritsky Institute of Geology and Geochemistry, Ural Branch of the Russian Academy of Sciences
(IGG UB RAS) 620110, Yekaterinburg, Akademika Vonsovskogo St., 15;
Kazan (Volga region) Federal University, Kazan University, KFU 420008, Russia, Republic of
Tatarstan, Kazan, 18 Kremlyovskaya street

INVESTIGATION OF COMPUTER VISION OBJECT DETECTION APPROACHES IN ROCK FRAGMENTATION PROBLEMS

В статье рассмотрена задача оценки фрагментации обломков горных пород на карьерах при помощи методов компьютерного зрения. Показаны преимущества использования современных одноэтапных архитектур обнаружения объектов типа YOLO в режимах обнаружения объектов и обнаружения ориентированных объектов по сравнению с традиционно используемыми подходами на основе экзemplарной сегментации типа Mask R-CNN. Также показано, что при правильном выборе порога уверенности в обнаружении качество работы архитектур по метрике mAP статистически не различимо. Порог может быть выбран путем минимизации расстояния Васерштейна между разметкой набора данных и полученными значениями по предложенному алгоритму.

Ключевые слова: фрагментация горных пород, компьютерное зрение, глубокие нейронные сети, гранулометрический состав, обнаружение объектов.

The rock fragmentation problem in open pits is explored using computer vision techniques. The benefits of employing modern, YOLO-based object detection architectures for both object detection and oriented object detection are demonstrated, in comparison to traditionally used methods based on Mask-R-CNN instance segmentation. It is shown that, with the appropriate selection of a confidence threshold for detection, the performance by the mean average precision (mAP) metric, is statistically indistinguishable. The threshold may be selected by minimizing the Wasserstein distance between the annotated and the obtained values using the proposed algorithm.

Key words: rock fragmentation, computer vision, deep neural networks, object detection, blast quality estimation.

Введение

Фрагментация горной массы – один из ключевых показателей эффективности буровзрывных работ при открытой добыче полезных ископаемых. На большинстве горных предприятий фрагментация взорванной горной массы производится вручную, посредством отбора изображений развалов взорванной горной массы, с последующим анализом, сочетающим ручную и автоматическую обработку изображений. Однако такой подход имеет существенные недостатки: он трудоёмок, экономически затратен и, что особенно важно, не обеспечивает полноты информации – например, о пространственном распределении фрагментов различных размеров по объёму развала или о локальных изменениях гранулометрического состава в пределах взрывного блока [1]. Это обуславливает актуальность задачи автоматизации контроля фрагментации горной массы [2], [3].

На сегодняшний день одним из наиболее популярных методов оценки размеров фрагментов горной породы является метод на основе систем компьютерного зрения [4-7]. Метод включает такие операции, как получение изображения участка, обнаружение фрагментов горной породы, оценка выбранных геометрических характеристик каждого фрагмента и оценка соответствующего распределения [5], [8].

Среди методов компьютерного зрения для анализа фрагментации можно отдельно выделить направление, использующее нейронные сети глубокого обучения [6], [9]. Этот подход позволяет решать задачи “end-to-end”, то есть без использования дополнительных операций, например, предобработки или постобработки изображений. Кроме того, преимуществом такого подхода является единство фреймворка работы с моделью и высокая способность к параллелизму, характерная для нейронных сетей [10]. В процессе эксплуатации к моделям на основе нейронных сетей часто предъявляются требования к оперативности работы и возможности запуска моделей на низкопроизводительных мобильных устройствах.

В рамках применения нейронных сетей задача оценки фрагментации горной породы часто сводится к экземплярной сегментации (instance segmentation). То есть к задаче поиска и сегментирования отдельных объектов на изображениях. Однако такой подход зачастую является избыточным. Он требует значительных вычислительных ресурсов и сложной постобработки, особенно при работе с широкоформатными изображениями с высокой плотностью объектов [11]. В ряде работ до сих пор используются проверенные, но относительно медленные архитектуры, такие как Mask R-CNN [2], [6], [12].

В последние годы наблюдается переход к более лёгким и быстрым решениям, основанным на обнаружении объектов (object detection) вместо экземплярной сегментации. В частности, широкое распространение получили архитектуры семейства YOLO (You Only Look Once), обеспечивающие обработку в реальном времени [13-16]. Современные версии, такие как YOLOv8, а также их расширенные модификации поддерживают не только детекцию, но и дополнительные режимы работы: экземплярную сегментацию (instance segmentation), обнаружение ориентированных ограничивающих рамок (oriented bounding boxes, OBB), учитывающих наклон и вытянутость фрагментов, предсказание ключевых точек (keypoint detection) [16-21]. Эти расширенные возможности позволяют более точно моделировать геометрию фрагментов, улучшая оценку их линейных размеров, формы и пространственной ориентации. При этом современные версии YOLO оптимизированы для развёртывания на ресурсоограниченных устройствах (включая мобильные и embedded-платформы) и способны обеспечивать инференс в реальном времени даже при ограниченных вычислительных мощностях [18]. При этом указанные дополнительные режимы работы могут представлять интерес в целях повышения точности решения задач фрагментации.

Цель работы – исследование современных архитектур обнаружения объектов и их модификаций для решения задач, связанных с оценкой фрагментации горных пород в условиях карьера в режиме реального времени. Задача формулируется как восстановление распределения размеров фрагментов пород на цифровом изображении с использованием четырёх режимов работы модели (обнаружение объектов, обнаружение ориентированных объектов, сегментация, ключевые точки) и калибровки порога уверенности путём минимизации расстояния Васерштейна. Вклад настоящей работы в область исследований заключается в том, что показаны преимущества решения задач фрагментации горных пород на карьерах больших масштабов как модификации задач компьютерного зрения: обнаружения объектов, обнаружение ключевых точек, сегментация объектов и обнаружение ориентированных объектов для архитектур на базе YOLOv8 в реальном масштабе времени.

Материалы и методы

Среди архитектур решения задач обнаружения объектов в реальном масштабе времени наиболее популярны и продуктивны архитектуры семейства YOLO [16]. Подход YOLO основан на идее, что входное изображение делится на ячейки одинакового размера, каждая из которых соответствует одному выходному вектору. Такой вектор может включать в себя одну или несколько привязок (анхоров, anchors) с предсказаниями размера ограничивающей рамки, оценкой класса и оценкой объектности. Развитие семейства архитектур YOLO производилось путем экспериментов с архитектурой и особенностями тренировки, как в области кодировщика признаков, так и в области головной и шейной частей. Индустриальный стандарт архитектур YOLO поддерживает компания Ultralytics [22]. Среди последних архитектур YOLO наиболее востребована модификация YOLOv8 [23]. Эта архитектура использует идею пирамидального извлечения карт признаков с агрегированием признаков, полученных с разных уровней (PAN). Агрегация признаков происходит с разных уровней кодировщика признаков путем передачи через остаточные связи. Таким образом, сеть может лучше улавливать признаки в различных масштабах и разрешениях, что очень важно для точного обнаружения объектов различных размеров и форм. В головной части использован подход без привязок (без т.н. анхоров) и две независимые параллельные головные части для оценки присутствия объекта (задача классификации) и регрессии обрамляющей рамки [18], [24].

Модель YOLOv8 имеет пять масштабированных конфигураций: YOLOv8n (нано), YOLOv8s (маленький), YOLOv8m (средний), YOLOv8l (большой) и YOLOv8x (экстремально большой). Архитектура YOLOv8 позволяет проводить определение объектов на изображении в режимах обнаружения объектов, обнаружения ориентированных объектов, поиск ключевых точек и экземплярной сегментации объектов. Иллюстрации интерпретации указанных подходов в отношении задачи фрагментации для отдельного фрагмента горной породы приведены в соответствии с рисунком 1 а) и б).

В задаче «обнаружение объектов» размеры обрамляющей рамки определяются через ширину w и высоту h . В качестве максимального размера фрагмента горной породы примем значение максимальной стороны рамки, описанной вокруг фрагмента горной породы (метод описанного прямоугольника) d_{max} . В случае подхода на основе ориентированной обрамляющей рамки угол наклона рамки ориентирован по направлению максимального размера d_{max} . В случае задачи обнаружения ключевых точек заданы четыре точки в соответствии с рисунком 1б. Точки выбраны так, что две ключевые точки, определяют максимальный размер внутри породы, а вторые две точки образуют линию, перпендикулярную первым двум.

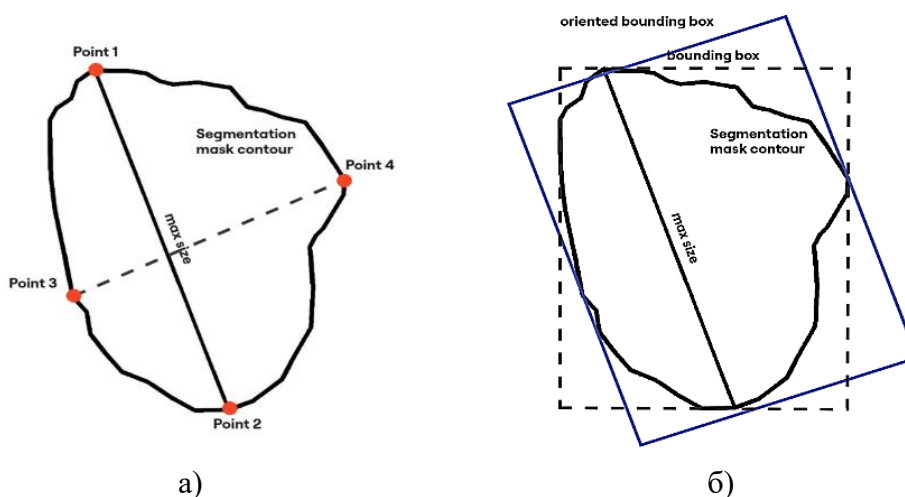


Рисунок 1 – Иллюстрации подходов к обнаружению объектов: а) обрамляющая рамка и ориентированная рамка; б) ключевые точки; сегментация предполагает определение истинных границ фрагмента

В проведённом исследовании использован собранный авторами набор данных, включающий 260 цифровых изображений участков карьера после буровзрывных работ [25], [26]. Съёмка выполнена на Баженовском месторождении (Свердловская область, г. Асбест). Основная часть изображений получена в ходе пяти полевых экспериментов, проведённых в летний и осенний периоды. В каждом эксперименте участвовало от 3 до 5 разрабатываемых участков, отобранных специалистами геологической службы как наиболее репрезентативные с точки зрения геологического строения и условий взрывания. Для каждого участка карьера было сделано несколько снимков разных его частей. Каждое изображение содержит порядка 20 – 80 фрагментов горных пород разного размера. Для всех изображений выполнена ручная разметка с использованием инструмента CVAT (Computer Vision Annotation Tool). Также в качестве аннотаций применены ориентированные обрамляющие рамки, выровненные вдоль направления максимального линейного размера каждого фрагмента. Пример изображения и его аннотация с помощью ориентированных обрамляющих рамок показана в соответствии с рисунком 2.

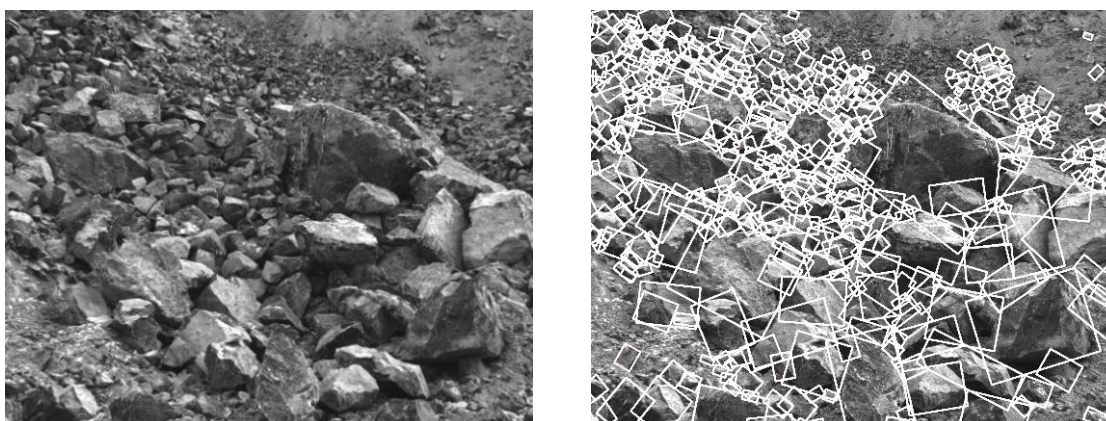


Рисунок 2 – Типичный пример изображений с обломками горной породы в карьере после взрывных работ: а) хризотил-асбестовые руды в серпентинитах, Свердловская область, г. Асбест; б) Разметка с помощью ориентированных обрамляющих рамок

Метрика оценивания качества фрагментации горных пород

Для оценки качества систем компьютерного зрения в задачах обнаружения и сегментации объектов традиционно применяются метрики, основанные на геометрическом совпадении предсказанных и размеченных объектов. Наиболее распространёнными из них являются:

- Intersection over Union (IoU) – отношение площади пересечения предсказанной и истинной областей к площади их объединения;
- mean Average Precision (mAP) – усреднённая точность обнаружения, рассчитываемая по кривой precision–recall, часто при пороге $\text{IoU} \geq 0.5$ (обозначается как $\text{mAP}@0.5$).

Однако эти метрики оценивают локальное качество обнаружения отдельных объектов и не отражают точность восстановления глобального распределения размеров фрагментов, что является главной целью в задачах фрагментационного анализа. В связи с этим предложено использовать расстояние Васерштейна (Wasserstein distance) [27]. Напомним, что расстояние Васерштейна можно интерпретировать как минимальные затраты энергии, которые необходимо совершить, чтобы переместить массу из одного распределения, чтобы заполнить другое распределение. Расстояние Васерштейна для распределений μ и η может быть задано как:

$$d_w(\mu, \eta) = \left(\int_0^1 |F^{-1}(x) - G^{-1}(x)|^2 dx \right)^{1/2},$$

где F^{-1} и G^{-1} квантильные функции распределений μ и η . Таким образом, расстояние Васерштейна является мерой схожести распределений предсказанных максимальных размеров фрагментов горной породы с размеченными максимальными размерами фрагментов. Для реализации оценки распределения с использованием расстояния Васерштейна реализован следующий алгоритм:

1. Выполняется прогнозирование обрамляющих рамок обученной моделью.
2. Выбираются обрамляющие рамки, для которых значение уверенности сети (confidence threshold) выше заданного порога уверенности. Формируется выборка значений размера обломков породы $R = \bigcup_{i=1}^K d_{\max_i}$, где K – количество предсказанных обломков породы.
3. Формируется выборка значений d_{\max} обломков породы на основе разметки $L = \bigcup_{i=1}^N l_i$.
4. Выполняется вычисление расстояния Васерштейна между выборками L и R с помощью (1). Отметим, что для вычисления расстояния Васерштейна используется функция `stats.wasserstein_distance` из математического пакета `scipy` [27].
5. Проводится анализ зависимости расстояния Васерштейна от уровня уверенности сети. Для этого выполняется варьирование порога уверенности сети от 0,05 до 0,75 с шагом 0,05 и вычисляется расстояние Васерштейна для распределений полученных на валидационном наборе изображений.
6. Выбирается такое значение порога, при котором достигается наименьшее значение расстояния Васерштейна.
7. Выбранное значение порога используется для обработки тестовых данных.

Для обучения, валидации и тестирования модели исходный набор изображений был разделён на три непересекающиеся подвыборки: обучающая выборка 170 изображений, валидационная выборка 25, тестовая выборка 65. В тестовую выборку вошли наиболее репрезентативные изображения разрабатываемых участков карьера. Для обучения моделей YOLO использовался фреймворк Ultralytics, позволяющий использовать разные режимы работы: экземплярная сегментация, обрамляющие рамки, ориентированные обрамляющие рамки и ключевые точки [22]. Все модели для разных режимов работы использовались в своих максимальных конфигурациях YOLOv8x. При проведении экспериментов использовался размер изображений 640 на 640 пикселей. Также в качестве базовой модели (т.н. baseline) использовалась модель Mask R-CNN. Модели YOLO и Mask R-CNN имели предобученные весовые параметры, полученные на наборе данных Common Objects in Context (COCO). Обучение моделей семейства YOLO проводилось со стандартными параметрами рекомендуемыми разработчиками [22]. Параметр случайного поворота изображения изменен на $\text{degrees}=45$. Для обучения установлено значение lr 0,01 для оптимизатора SGD, число эпох обучения 100, размер батча 4.

Результаты экспериментов и их анализ

В таблице 1 показаны результаты метрик полученных для тестового набора данных, где: N - количество обнаруженных фрагментов, интерполированная средняя точность (mean Average Precision) для порогового значения IOU 0,5 ($\text{mAP}@0.5$); значение IOU, расстояние Васерштейна ($d_{0,25}$) при заданном пороге уверенности сети в 0,25; время обработки одного изображения в миллисекундах (t_{cpu}). Значение IOU и mAP приведено для справки ввиду невозможности корректного сравнения предсказаний между собой для ориентированных обрамляющих рамок и экземплярной сегментации.

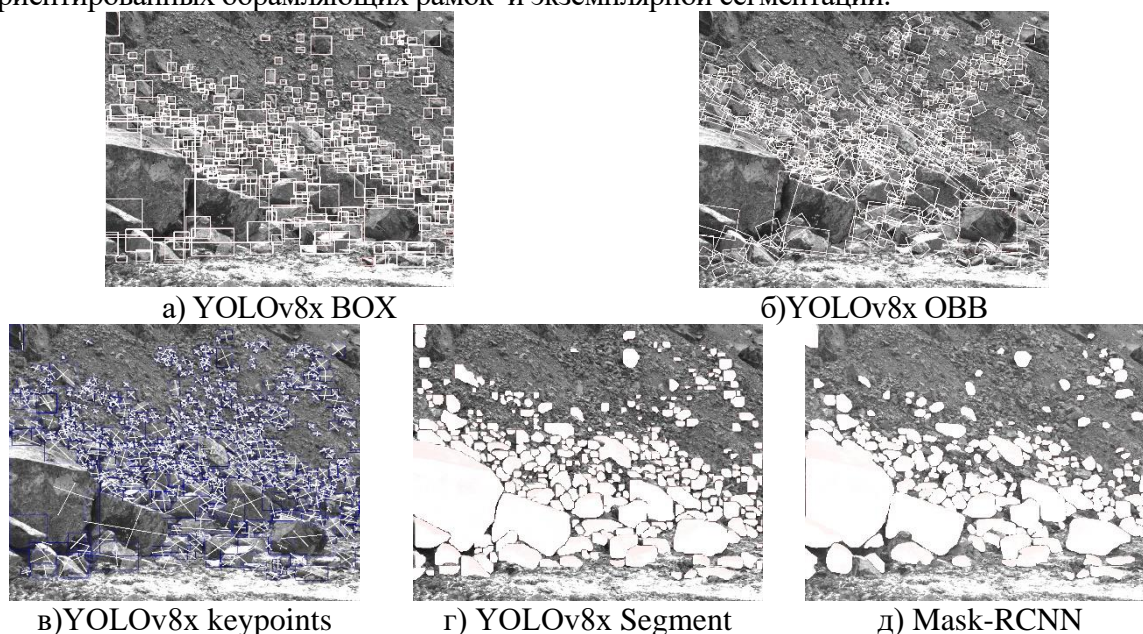


Рисунок 3 – Типичные визуализации результатов обнаружения фрагментов разрушенной горной породы для тестовых изображений: а – обнаружение объектов; б – обнаружение ориентированных объектов; в – обнаружение ключевых точек; г – объектная сегментация; д – объектная сегментация Mask-RCNN

Оценка времени обработки одного изображения выполнялась на процессоре Intel Core(TM) i5-11400H 2,7 ГГц (CPU) и на видеокарте Tesla V100-SXM2 (GPU). В таблице приняты следующие обозначения режимов работы моделей: BOX - задача обнаружения объектов; OBB - обнаружение ориентированных обрамляющих рамок; Seg - экземплярная сегментация; KP - ключевые точки. Типичные визуализации результатов обнаружения фрагментов разрушенной горной породы для тестовых изображений показаны на рисунках 3 а)–д) для исследуемых типов задач. Отметим, что на рисунке 3 в) для каждого фрагмента горной породы ключевые точки, определяющие два размера фрагмента, соединены с помощью линий. Можно отметить, что оценка размеров через ключевые точки систематически занижает истинный размер фрагмента породы.

Таблица 1 – Метрики обученных моделей на тестовом наборе изображений

Модель	N	mAP@0.5	IOU	$d_{0,25}$	$t_{gpu}, \text{мс}$	$t_{inf CPU}, \text{мс}$
YOLOv8x BOX	16564	0,68	0,83	0,0042	30,1	406
YOLOv8x OBB	18098	0,67	0,81	0,0041	30,4	416
YOLOv8x KP	15147	0,61	0,83	0,6442	31,4	469
YOLOv8x Seg	14113	0,64	0,73	0,0039	35,9	569
Mask-RCNN cascade 3x	12688	0,58	0,71	0,0022	172,1	2427

Из таблицы 1 видно, что наилучшее значение метрик по точности mAP для моделей с обрамляющими рамками демонстрирует YOLOv8x в режимах обнаружения объектов и сегментации объектов. При этом экземплярная сегментация имеет меньшую точность как для моделей YOLOv8 так и для базовой модели Mask-RCNN cascade 3x. Из рисунка 3 видно, что модель Mask-RCNN cascade 3x сегментирует меньшее число фрагментов разрушенной горной породы по сравнению с другими моделями. Также модель Mask-RCNN cascade 3x уступает по скорости обработки моделям YOLO в 6 - 7 раз, что ограничивает возможность практического использования Mask-RCNN в режимах реального времени.

Анализ расстояния Васерштейна между распределениями

Была проанализирована зависимость расстояния Васерштейна от значения порога уверенности модели (confidence threshold). Порог уверенности определяет минимальную вероятность, при которой предсказанная моделью локальная гипотеза (bounding box) принимается в качестве достоверного обнаружения целевого объекта. Порог уверенности сети изменялся в диапазоне значений от 0,1 до 0,75 с шагом 0,05. Для каждого значения параметра было подсчитано расстояние Васерштейна между оцененным распределением размеров фрагментов и распределением размеров из разметки для тестового набора изображений. Зависимость значений расстояния Васерштейна от выбора порога уверенности показаны на рисунке 4.

Из анализа рисунка 4 видно, что минимальное расстояние Васерштейна достигается при пороге уверенности сети 0,25 для всех моделей. При выбранном пороге значение метрики mAP@0.5 составили: Seg(0,65), BOX(0,67), OBB(0,69), а для Mask-RCNN mAP@0.5 (0,58). Также можно отметить, что результаты архитектуры Mask-RCNN более устойчивы к выбору порога уверенности, чем для архитектуры YOLO.

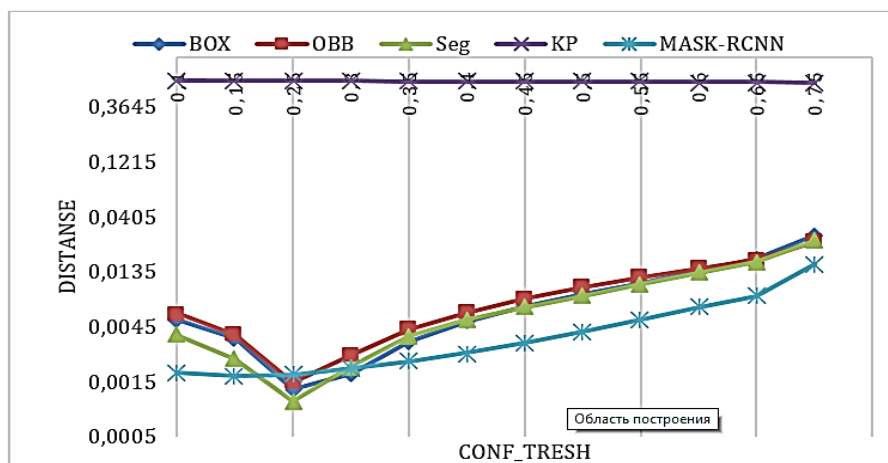


Рисунок 4 – Типичные графики значений расстояния Васерштейна между оцененным распределением размеров фрагментов и распределением размеров из разметки для тестового набора данных

В силу «на глаз» схожих результатов для подходов BOX и OBB была проверена гипотеза об отсутствии различия двух выборок согласно U-критерию Манна - Уитни (гипотеза H_0). Использовалась функция `stats.mannwhitneyu` из пакета `scipy`. Указанная гипотеза подтверждается на уровне значимости 0,05, значение статистики 138876348,0 и $p_value=0,957$. Таким образом, можно утверждать об отсутствии статистического различия в подходах обрамляющих рамок (BOX) и ориентированных обрамляющих рамок (OBB).

Обсуждение результатов

Результаты, представленные на рисунках 3, 4 и в таблице 1, убедительно демонстрируют возможность эффективного использования моделей семейства YOLO в решении задач фрагментации горных пород. Проведенный анализ показывает, что данные модели способны обеспечивать сопоставимое с другими подходами качество восстановления распределений размеров фрагментов на изображениях. Результаты для моделей YOLO дают сопоставимые значения для исследованных подходов на основе обнаружения объектов, обнаружения ориентированных объектов и сегментации объектов. Это может быть объяснено достаточно большим объемом фрагментов, представленных на изображениях. При этом часто фрагменты ориентированы в схожих направлениях и имеют выпуклые геометрические формы. Таким образом, предпочтение может быть отдано более быстрым методам, основанным на обнаружении объектов в стандартных или ориентированных ограничивающих рамках. При этом статистически значимой разницы в точности между подходами с использованием ориентированных и неориентированных рамок выявлено не было.

Сравнение результатов для моделей YOLO с базовым решением на основе архитектуры Mask-RCNN показывает преимущества последней в точности восстановления распределений. Однако, ключевым преимуществом архитектур YOLO является их производительность. Скорость обработки изображений на GPU для моделей YOLOv8 более чем в 5 раз превышает скорость Mask R-CNN (30-36 мс против 172 мс), что делает их оптимальным выбором для систем, требующих анализа в режиме реального времени. Несмотря на то, что Mask R-CNN демонстрирует наилучшую точность восстановления распределения размеров (минимальное расстояние Васерштейна (0,0022), модели YOLOv8 обеспечивают очень близкое качество (0,0039–0,0042) при значительно меньших вычислительных затратах.

Выводы

В настоящей работе проведено исследование современных архитектур компьютерного зрения для решения задачи оценки фрагментации горных пород в карьерных условиях в режиме реального времени. Показано, что задача оценки фрагментации может быть эффективно решена с использованием архитектуры YOLOv8 в рамках нескольких подходов: обнаружение объектов (object detection), экземплярная сегментация (instance segmentation), обнаружение ориентированных ограничивающих рамок (oriented bounding boxes) и распознавание ключевых точек (keypoint detection). Результаты показывают качество восстановления распределений размеров фрагментов горной породы, сопоставимое с базовой архитектурой Mask-RCNN при калибровке (выборе) порогового значения уверенности сети. Однако, при этом архитектуры на базе YOLOv8 имеют скорость вычислений в 5 раз выше и позволяют проводить измерения на центральных процессорах за доли секунды.

Выбор порога уверенности сети (калибровка) может быть выполнен при помощи минимизации расстояния Васерштейна согласно описанному алгоритму. При этом значения метрик mAP@0.5 равны Seg(0,65), BOx(0,67), OBB(0,69), а для Mask-RCNN(0,58). Результаты для подходов обнаружения объектов и обнаружения ориентированных объектов не имеют статистических различий. Следует также отметить качественное соответствие подхода обнаружения ориентированных объектов интуитивному пониманию особенностей работы этих алгоритмов. Поэтому данный подход может быть рекомендован как основной в дальнейших исследованиях.

Список литературы

1. Дробление горных пород взрывом в карьерах / Н. Н. Казаков, С. Д. Викторов, А. В. Шляпин, И. Н. Лапиков ; Рос. акад. наук. Москва : Наука, 2020. 519 с.: ил., табл. Библиогр.: с. 509–518. Прил.: с. 418–508. ISBN 978-5-907036-96-3.
2. Vu, T. Measuring blast fragmentation at Nui Phao open-pit mine, Vietnam using the Mask R-CNN deep learning model / T. Vu, T. Bao, Q. V. Hoang [et al.] // Mining Technology. 2021. Vol. 130, No. 4. P. 232–243. DOI 10.1080/25726668.2021.1944458.
3. Оптимизация фрагментации взорванной горной массы [Электронный ресурс] // Геология и бизнес:[сайт]. [Б. м.], [2023]. URL: <https://www.vnedra.ru/tehnologii/optimizacziya-fragmentaczii-vzorvannoj-gornoj-massy-8638> (дата обращения: 19.12.2025).
4. Williams, J. Exploring Deep Learning for Dig-Limit Optimization in Open-Pit Mines / J. Williams, J. Singh, M. Kumral, J. Ramirez-Ruiseco // Natural Resources Research. 2021. Vol. 30, No. 3. P. 2085–2101. DOI 10.1007/s11053-021-09864-y.
5. Mkwelo, S. Watershed-based segmentation of rock scenes and Proximity-based classification of watershed regions under uncontrolled lighting conditions [Электронный ресурс] / S. Mkwelo, G. de Jager, F. Nicolls. – 2003. – URL: <https://api.semanticscholar.org/CorpusID:9527232> (дата обращения: 19.12.2025).
6. Bamford, T. A deep learning approach for rock fragmentation analysis / T. Bamford, K. Esmaeili, A. P. Schoellig // International Journal of Rock Mechanics and Mining Sciences. 2021. Vol. 145. Art. 104839.
7. Peng, J. Ore Rock Fragmentation Calculation Based on Multi-Modal Fusion of Point Clouds and Images / J. Peng [et al.] // Applied Sciences. 2023. Vol. 13, No. 23. Art. 13542.
8. Guo, Q. A method of blasted rock image segmentation based on improved watershed algorithm / Q. Guo [et al.] // Scientific Reports. 2022. Vol. 12, No. 1. Art. 7143.
9. Каспаров, Ю. А. Обучение моделей компьютерного зрения для задачи детекции и сегментации стволовых клеток / Ю. А. Каспаров, В. М. Зуев // Проблемы искусственного интеллекта. 2025. № 3 (38). С. 183–189. DOI 10.24412/2413-7383-2025-3-38-183-189.
10. Jia, Z. Beyond Data and Model Parallelism for Deep Neural Networks / Z. Jia, M. Zaharia, A. Aiken // Proceedings of Machine Learning and Systems (MLSys 2019), Stanford, CA, USA, March 31 – April 2, 2019 / ed. by A. Talwalkar, V. Smith, M. Zaharia. [S. l.], 2019. Vol. 1. P. 1–13.
11. Tang, X. A new ore image segmentation method based on Swin-UNET / X. Tang [et al.] // 2022 China Automation Congress (CAC): proceedings. [S. l.]: IEEE, 2022. P. 1681–1686.

12. Tang, W. A remote sensing image rotation object detection approach for real-time environmental monitoring / W. Tang [et al.] // *Sustainable Energy Technologies and Assessments*. 2023. Vol. 57. Art. 103270.
13. Tao, M. A Fast Recognition Method for Dynamic Blasting Fragmentation Based on YOLOv8 and Binocular Vision / M. Tao [et al.] // *Applied Sciences*. 2025. Vol. 15, No. 12. Art. 6411.
14. Wang, Q. Advances in Deep Learning-Based Ore Particle Size Detection: A Review of Methods, Challenges, and Trends / Q. Wang [et al.] // *MetaResource*. 2025. Vol. 2, No. 2. P. 83–104.
15. Wang, X. An Appropriate Approach to Recognize Coke Size Distribution in a Blast Furnace / X. Wang, Y. Guo, Y. Yu // *Processes*. – 2023. – Vol. 11, No. 1. – Art. 187.
16. Redmon, J. YOLOv3: An Incremental Improvement [Электронный ресурс] / J. Redmon, A. Farhadi // arXiv preprint. – 2018. – URL: <https://arxiv.org/abs/1804.02767> (дата обращения: 19.12.2025).
17. Shrivastava, S. Segmentation of mine overburden dump particles from images using Mask R-CNN / S. Shrivastava, S. Bhattacharjee, D. Deb // *Scientific Reports*. 2023. Vol. 13, No. 1. Art. 2046. DOI 10.1038/s41598-023-28586-0.
18. Terven, J. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS / J. Terven, D. M. Córdova-Esparza, J. A. Romero-González // *Machine Learning and Knowledge Extraction*. 2023. Vol. 5, No. 4. P. 1680–1716. DOI 10.3390/make5040083.
19. Sun, Z. GOG-RT-DETR: An Improved RT-DETR-Based Method for Graphite Ore Grade Detection / Z. Sun [et al.] // *Applied Sciences*. 2025 Vol. 15, No. 24. Art. 14287.
20. Meng, H. Improved YOLOv11-EGM deep learning model for rock fragment identification / H. Meng [et al.] // *Archives of Civil and Mechanical Engineering*. 2025. Vol. 25, No. 4. Art. 223.
21. Sabharwal, M. Real-time detection of coke particles in blast furnace operations using machine learning: Case of a steel plant in India / M. Sabharwal [et al.] // *Ironmaking & Steelmaking*. 2024. Vol. 52, No. 4. P. 412–425.
22. YOLOv8 [Электронный ресурс] // *Ultralytics*. 2024. URL: <https://docs.ultralytics.com/> (дата обращения: 19.12.2025).
23. Redmon, J. YOLO9000: Better, Faster, Stronger / J. Redmon, A. Farhadi // *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI, USA : IEEE, 2017. P. 6517–6525. DOI 10.1109/CVPR.2017.690.
24. Ronkin, M. Real-Time YOLO-family Comparison for Blast Quality Estimation in the Open Pit Conditions / M. Ronkin, K. Reshetnikov // *Proceedings of 2023 IEEE Ural-Siberian Conference on Biomedical Engineering, Radioelectronics and Information Technology (USBREIT)*. Piscataway : IEEE, 2023. P. 254–257.
25. Решетников, К. И. Исследование подхода «обнаружение объектов» в задаче фрагментации горных пород на открытых карьерах / К. И. Решетников, М. В. Ронкин, С. В. Поршнева // *Компьютерная оптика*. 2024. Т. 48, № 2. С. 272–281. DOI 10.18287/2412-6179-CO-1382.
26. Ronkin, M. openpits asbestos [Электронный ресурс] / M. Ronkin, K. Reshetnikov, V. Zyuzin // *Mendeley Data*. 2024, DOI 10.17632/pfdbfpygh.3 URL: <https://data.mendeley.com/datasets/pfdbfpygh/3> (дата обращения: 19.12.2025).
27. SciPy 1.0: fundamental algorithms for scientific computing in Python / P. Virtanen [et al.] // *Nature Methods*. 2020. Vol. 17, No. 3. P. 261–272.

References

1. Kazakov NN, Viktorov SD, Shlyapin AV, Lapikov IN. *Crushing of rocks by explosion in open pits*. Moscow: RAS Publ.; 2020. 519 p. ISBN 978-5-907036-96-3.
2. Vu T, Bao T, Hoang QV, Drebenstedt C, Hoa PV, Thang HH. Measuring blast fragmentation at Nui Phao open-pit mine, Vietnam using the Mask R-CNN deep learning model. *Mining Technology*. 2021;130(4):232–243. doi: 10.1080/25726668.2021.1944458.
3. Optimization of blasted rock mass fragmentation. *Geology and Business*. [онлайн]. URL: <https://www.vnedra.ru/tehnologii/optimizacziya-fragmentaczii-vzorvannoj-gornoj-massy-8638> (Дата доступа 19, 2025).
4. Williams J, Singh J, Kumral M, Ramirez-Ruiseco J. Exploring deep learning for dig-limit optimization in open-pit mines. *Natural Resources Research*. 2021;30(3):2085–2101. doi: 10.1007/s11053-021-09864-y.
5. Mkwelo S, de Jager G, Nicolls F. Watershed-based segmentation of rock scenes and proximity-based classification of watershed regions under uncontrolled lighting conditions. [Online]. Available at: <https://api.semanticscholar.org/CorpusID:9527232> (accessed December 19, 2025).
6. Bamford T, Esmaeili K, Schoellig A. A deep learning approach for rock fragmentation analysis. *International Journal of Rock Mechanics and Mining Sciences*. 2021;145:104839. doi: 10.1016/j.ijrmmms.2021.104839.

7. Peng J, et al. Ore rock fragmentation calculation based on multi-modal fusion of point clouds and images. *Applied Sciences*. 2023;13(23):12845. doi: 10.3390/app132312845.
8. Guo Q, et al. A method of blasted rock image segmentation based on improved watershed algorithm. *Scientific Reports*. 2022;12:7143. doi: 10.1038/s41598-022-09143-8.
9. Kasparov YA, Zuev VM. Training computer vision models for stem cell detection and segmentation task. *Problems of Artificial Intelligence*. 2025;3(38):183–189. ISSN 2413-7383. doi: 10.24412/2413-7383-2025-3-38-183-189.
10. Jia Z, Zaharia M, Aiken A. Beyond data and model parallelism for deep neural networks. In: Talwalkar A, Smith V, Zaharia M, editors. *Proceedings of Machine Learning and Systems*. Vol. 1. Stanford, CA, USA: MLSys; 2019. p. 1–13.
11. Tang X, et al. A new ore image segmentation method based on Swin-Unet. In: 2022 China Automation Congress (CAC). IEEE; 2022. p. 1681–1686. doi: 10.1109/CAC56365.2022.10075678.
12. Tang W, et al. A remote sensing image rotation object detection approach for real-time environmental monitoring. *Sustainable Energy Technologies and Assessments*. 2023;57:103270. doi: 10.1016/j.seta.2023.103270.
13. Tao M, et al. A fast recognition method for dynamic blasting fragmentation based on YOLOv8 and binocular vision. *Applied Sciences*. 2025;15(12):6411. doi: 10.3390/app15126411.
14. Wang Q, et al. Advances in deep learning-based ore particle size detection: A review of methods, challenges, and trends. *MetaResource*. 2025;2(2):83–104. doi: 10.3390/metaresource2020008.
15. Wang X, Guo Y, Yu Y. An appropriate approach to recognize coke size distribution in a blast furnace. *Processes*. 2023;11(2):187. doi: 10.3390/pr11020187.
16. Redmon J, Farhadi A. YOLOv3: An incremental improvement. arXiv preprint arXiv:1804.02767. 2018.
17. Shrivastava S, Bhattacharjee S, Deb D. Segmentation of mine overburden dump particles from images using Mask R-CNN. *Scientific Reports*. 2023;13:2046. doi: 10.1038/s41598-023-28586-0.
18. Terven J, Córdova-Esparza DM, Romero-González JA. A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS. *Machine Learning and Knowledge Extraction*. 2023;5(4):1680–1716. doi: 10.3390/make5040083.
19. Sun Z, et al. GOG-RT-DETR: An improved RT-DETR-based method for graphite ore grade detection. *Applied Sciences*. 2025;15(24):14567. doi: 10.3390/app152414567.
20. Meng H, et al. Improved YOLOv11-EGM deep learning model for rock fragment identification. *Archives of Civil and Mechanical Engineering*. 2025;25(4):223. doi: 10.1007/s43452-025-01234-5.
21. Sabharwal M, et al. Real-time detection of coke particles in blast furnace operations using machine learning: Case of a steel plant in India. *Ironmaking & Steelmaking. Processes, Products and Applications*, vol. 52, Oct. 2024, doi: 10.1177/03019233241276285.
22. YOLOv8 documentation. Ultralytics Inc. [Online]. Available at: <https://docs.ultralytics.com/> (accessed December 19, 2025).
23. Redmon J, Farhadi A. YOLO9000: Better, faster, stronger. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE; 2017. p. 6517–6525. doi: 10.1109/CVPR.2017.690.
24. Ronkin M, Reshetnikov K. Real-time YOLO-family comparison for blast quality estimation in the open pit conditions. In: 2023 IEEE Ural-Siberian Conference on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT). Piscataway, NJ, USA: IEEE; 2023. p. 254–257. doi: 10.1109/USBEREIT58456.2023.10284567.
25. Reshetnikov KI, Ronkin MV, Porshnev SV. Investigation of the object detection approach in the rock fragmentation problem at open-pit mines. *Computer Optics*. 2024;48(2):272–281. doi: 10.18287/2412-6179-CO-1382.
26. Ronkin M, Reshetnikov K, Zyuzin V. Openpits asbestos dataset. *Mendeley Data*. V3. [Online]. Available at: <https://doi.org/10.17632/pfdbfpfygh.3> (accessed December 19, 2025).
27. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*. 2020;17(3):261–272. doi: 10.1038/s41586-019-1855-2.

RESUME

K. I. Reshetnikov, M. S. Glukhov, M. V. Ronkin

Investigation of Computer Vision Object Detection Approaches in Rock Fragmentation Problems

Automation of rock fragmentation analysis in open-pit mining is critical for optimizing blasting efficiency. The traditionally used computer vision methods, such as those based on the Mask R-CNN architecture, provide high accuracy but suffer from low processing speed, which makes them impractical for real-time applications. In this paper, we propose a method that combines the high performance of modern single-stage detectors like YOLO with a novel calibration technique to maintain high accuracy in estimating particle size distribution.

The study utilized the YOLOv8x architecture, evaluated in several modes including standard object detection (BOX), oriented bounding boxes (OBB), and instance segmentation, using a custom dataset of 260 images. The core of the proposed method is a calibration algorithm for selecting the optimal confidence threshold. The threshold is selected by systematically minimizing the Wasserstein distance between the predicted fragment size distribution and the ground-truth distribution from the annotated data.

The YOLOv8-based models demonstrated a processing speed more than 5 times faster than the Mask R-CNN baseline, enabling real-time analysis. While Mask R-CNN achieved a marginally better accuracy in distribution similarity (Wasserstein distance of 0.0022 vs. 0.0039–0.0042), the YOLO models yielded highly comparable results at a fraction of the computational cost. After calibration, the YOLO models (mAP up to 0.69) outperformed Mask R-CNN (mAP of 0.58) in object detection quality.

The developed approach, which combines the YOLOv8 architecture with a calibration method based on Wasserstein distance, is proven to be an effective and practical solution for real-time rock fragmentation analysis. It successfully brings together the high speed of single-stage detectors and the high accuracy of traditional methods, making it a superior choice for integration into industrial operational processes.

РЕЗЮМЕ

К. И. Решетников, М. С. Глухов, М. В. Ронкин

Исследование подходов компьютерного зрения для обнаружения объектов в задачах фрагментации обломков горных пород

В данной статье разработана и исследована методика оценки гранулометрического состава взорванной горной массы на базе современных архитектур компьютерного зрения.

На базе архитектуры YOLOv8 предложены и проанализированы подходы к решению задачи как обнаружения объектов, обнаружения ориентированных объектов, определение ключевых и экземплярной сегментации. Показано, что предложенные подходы обеспечивают высокую производительность, более чем в пять раз превышающую Mask R-CNN, при сохранении сопоставимого качества восстановления гранулометрического состава.

Разработан алгоритм калибровки моделей на основе минимизации расстояния Васерштейна, позволяющий достичь качества восстановления распределения размеров фрагментов, сопоставимого с более медленными аналогами.

Разработанная методика позволяет однозначно определить гранулометрический состав в режиме реального времени. Реализация системы с учётом предложенной методики позволяет решить задачу повышения эффективности контроля буровзрывных работ за счёт использования высокопроизводительных нейросетевых моделей.

Решетников Кирилл Игоревич – ассистент, учебно-научный центр «Искусственный интеллект» ФГАОУ ВО «УрФУ имени первого Президента России Б.Н. Ельцина», г. Екатеринбург. *Область научных интересов:* нейронные сети, машинное зрение, эл. почта ki.reshetnikov@urfu.ru, адрес 620002, г. Екатеринбург, ул. Мира, д. 19, телефон +79539136389, <https://orcid.org/0009-0006-1526-0368>.

Ронкин Михаил Владимирович – кандидат технических наук, доцент, учебно-научный центр «Искусственный интеллект» ФГАОУ ВО «УрФУ имени первого Президента России Б.Н. Ельцина», г. Екатеринбург. *Область научных интересов:* компьютерное зрение, методы обработки сигналов, машинное обучение, эл. почта m.v.ronkin@urfu.ru, адрес 620002, г. Екатеринбург, ул. Мира, д. 19, телефон +79222918776, <https://orcid.org/0000-0001-5522-988X>.

Глухов Михаил Сергеевич – кандидат географических наук, научный сотрудник Института геологии и геохимии имени акад. А.Н. Заварицкого Уральского отделения Российской академии наук (ИГГ УрО РАН) г. Екатеринбург; доцент Институт геологии и нефтегазовых технологий Казанского (Приволжского) федерального университета (К(П)ФУ) г. Казань. *Область научных интересов:* геохимия, палеогеология, прикладная минералогия, эл. почта Gluhov@igg.uran.ru, адрес 620002 Екатеринбург, ул. Академика Вонсовского, 15, телефон +79178763743, <https://orcid.org/0000-0002-5075-7066>.

Reshetnikov Kirill Igorevich – Assistant, educational and research Center "Artificial Intelligence", Ural Federal University named after the first President of Russia B.N. Yeltsin (UrFU), Yekaterinburg, Russia. Research interests: neural networks, computer vision. Email: ki.reshetnikov@urfu.ru. Address: 19 Mira Street, Yekaterinburg, 620002, Russia. Phone: +79539136389. ORCID: <https://orcid.org/0009-0006-1526-0368>.

Ronkin Mikhail Vladimirovich – Ph.D. in Engineering, Associate Professor, Educational and Research Center "Artificial Intelligence", Ural Federal University named after the first President of Russia B.N. Yeltsin (UrFU), Yekaterinburg, Russia. Research interests: computer vision, signal processing methods, machine learning. Email: m.v.ronkin@urfu.ru. Address: 19 Mira Street, Yekaterinburg, 620002, Russia. Phone: +79222918776. ORCID: <https://orcid.org/0000-0001-5522-988X>.

Glukhov Mikhail Sergeevich – Ph.D. in Geographical Sciences, Research Scientist, Zavaritsky Institute of Geology and Geochemistry, Ural Branch of the Russian Academy of Sciences (IGG UB RAS), Yekaterinburg, Russia; Associate Professor, Institute of Geology and Oil and Gas Technologies, Kazan (Volga Region) Federal University, KFU. Research interests: geochemistry, paleogeology, applied mineralogy. Email: Gluhov@igg.uran.ru. Address: 15 Akademika Vonsovskogo Street, Yekaterinburg, 620002, Russia. Phone: +79178763743. ORCID: <https://orcid.org/0000-0002-5075-7066>.

Статья поступила в редакцию 28.01.2026.