

Применение нейросетевой сегментации и оконтуривания объектов для оценки гранулометрического состава результата буровзрывных работ по изображению

В.Ю. Устимов, Д.А. Полещенко, И.С. Михайлов

*Старооскольский технологический институт им. А.А.Угарова (филиал)
Национального исследовательского технологического университета «МИСИС», Россия,
г. Старый Оскол*

Аннотация: Настоящее исследование посвящено разработке методов автоматической оценки гранулометрического состава руды после взрывных работ на основе данных, полученных с беспилотных летательных аппаратов (БПЛА). Определение размеров фрагментов руды является важным этапом, так как от его точности зависит эффективность последующих процессов дробления. В традиционных методах анализа используется ручная работа, что требует значительных трудозатрат и подвержено субъективным факторам. В исследовании рассматриваются современные методы машинного обучения и архитектуры нейронных сетей, такие как Feature Pyramid Network (FPN), EfficientNet и SE ResNet, которые могут автоматически и точно сегментировать изображения. В результате экспериментов было установлено, что наибольшую точность IoU среди моделей показала FPN-сеть с предобученным основанием EfficientNet B2.

Ключевые слова: гранулометрический состав, FPN, EfficientNet, SEResNet.

Введение

На сегодняшний день изучение гранулометрического состава горной породы после взрыва является важной задачей, поскольку от качества взрыва зависит эффективность всех последующих этапов переработки руды.

Оптимальный размер кусков, поступающих с рудника, составляет от 0 до 1200 мм, поскольку такие размеры может обрабатывать оборудование для крупного дробления. Наиболее эффективным считается диапазон 0–350 мм, так как такие фрагменты могут сразу направляться на следующий этап обогащения, минуя дробление.

Если после взрыва остаются куски руды диаметром более 2000 мм, для их дополнительного дробления используется гидравлический бутбой, что приводит к увеличению затрат. В связи с этим необходима новая система

оценки качества взрывных работ, которая позволит снизить расходы на дополнительные этапы дробления крупных фрагментов.

Цель данной работы – разработка подхода для оценки распределения размеров кусков руды после взрыва по данным фотографий с БПЛА, что позволит автоматизировать процесс анализа и повысить точность оценки эффективности взрывных работ.

1. Анализ существующих методов

Традиционные методы, такие как ручной отбор и физическое осаживание, требуют значительных человеческих ресурсов и подвержены субъективным факторам [1]. В последние годы были разработаны точные методы обработки изображений, которые позволяют автоматизировать процессы анализа и измерения размеров частиц. Эти методы включают в себя машинное обучение и алгоритмы глубокого обучения, способные эффективно сегментировать изображения руды и определять размеры частиц с высокой точностью [2]. Современные подходы обеспечивают объективность и устойчивость к внешним факторам, таким как освещение и текстура, что значительно улучшает качество получаемых данных и снижает потребность в человеческом участии.

В работе [3] рассматривается измерение размера частиц руды на конвейерной ленте. Представлена модель U-Net, которая использует многомасштабное соединение и учёт границ для более точной сегментации изображений. Для оценки размера руды по сегментированному изображению использован метод обнаружения дефектов с использованием алгоритма выпуклой оболочки (convex-hull defect detection, CDD). Полученный подход превосходит 24 алгоритма, включая классические методы обработки изображений и глубокое обучение.

Авторами исследования [4] предложена модель для сегментации частиц руды и фона FRRSnet+, которая имеет высокую точность ($R^2 = 99,58\%$) и малый размер.

2. Обзор архитектур нейронных сетей

Одним из критериев выбора структуры нейронной сети является наличие объектов различных размеров на изображении. Проблема наличия объектов разных размеров в изображении заключается в том, что нейронная сеть должна уметь эффективно обрабатывать как мелкие, так и крупные объекты. Мелкие объекты могут теряться на поздних этапах свёрточной обработки, когда разрешение изображений снижается, а крупные объекты могут оказаться недостаточно точно детализированы. Поэтому для надёжного распознавания всех объектов требуется архитектура, способная учитывать разный масштаб, как, например, Feature Pyramid Network (FPN) [5], которая объединяет признаки с разных уровней, сохраняя баланс между детализацией и обобщением.

Сеть FPN включает (см. рисунок 1):

- восходящий путь — иерархическая пирамида свёрточных слоёв, где изображение уменьшается, а обобщающая способность увеличивается, однако часть информации может теряться;
- нисходящий путь — структура, увеличивающая разрешение изображения для усиления семантики;
- боковые связи — помогают передать детальную информацию от исходного изображения.

Структура пирамиды представлена на рисунке 5.

В данной работе использована техника трансферного обучения с использованием предобученных оснований.

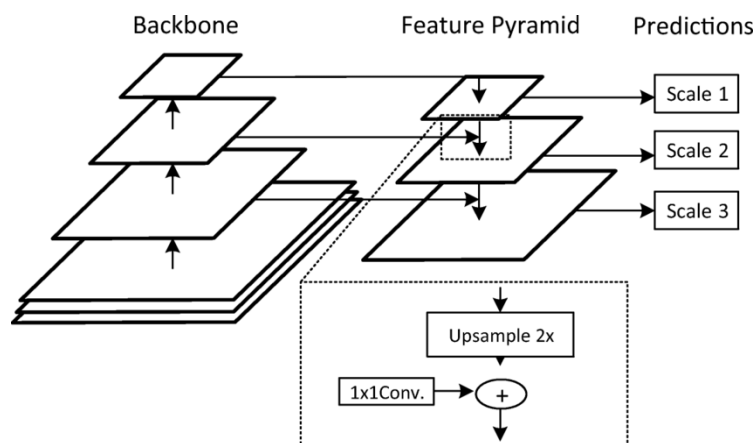


Рис. 1. – Архитектура FPN-сети [6]

EfficientNet — это семейство предобученных оснований НС, основанных на методе составного масштабирования (compound scaling method), который решает проблему баланса между размером модели, её точностью и вычислительной эффективностью. Эта концепция предполагает масштабирование трёх ключевых параметров сети: ширины, глубины и разрешения [7].

- **Ширина:** увеличение числа каналов в каждом слое позволяет модели выявлять более сложные паттерны и признаки, повышая точность, но увеличивает ресурсоёмкость. Уменьшение ширины делает модель легче, подходящей для сред с ограниченными ресурсами.
- **Глубина:** увеличение числа слоев позволяет модели глубже анализировать данные, но требует больше ресурсов. Меньшее число слоев экономит ресурсы, но может снизить точность.
- **Разрешение:** изменение разрешения входного изображения позволяет контролировать уровень деталей, которые модель может обработать. Более высокое разрешение повышает точность, но требует больше памяти и мощности. Низкое разрешение экономит ресурсы, но снижает детализацию.

Одним из достоинств EfficientNet является подход к согласованному масштабированию этих параметров через коэффициент составного

масштабирования (ϕ), который одновременно регулирует ширину, глубину и разрешение, помогая найти оптимальный баланс между точностью и сложностью модели.

SE ResNet — это разновидность нейронной сети ResNet [8], улучшенная за счёт использования модуля Squeeze-and-Excitation (SE), который добавляет внимание к значимости каналов [9]. Этот подход позволяет сети адаптивно усиливать важные признаки и подавлять менее значимые, что улучшает точность при небольшом увеличении вычислительных затрат.

Модуль Squeeze-and-Excitation включает два ключевых этапа:

- **Squeeze:** Пространственная агрегация информации, где модель сжимает данные по пространственным измерениям, сохраняя только информацию о каналах. Это уменьшает объём информации, сохраняя её ключевые черты.
- **Excitation:** на основе данных сжатия вычисляются весовые коэффициенты для каждого канала. Эти коэффициенты используются для акцентирования значимых каналов и уменьшения влияния менее важных.

SE ResNet сочетает мощь базовой архитектуры ResNet, которая эффективно решает проблемы исчезающего градиента и глубокой передачи признаков, с механизмом SE. Это улучшает её способность адаптироваться к различным признакам, повышая точность классификации и распознавания сложных объектов.

3. Подготовка данных

Экспериментальный участок результата буровзрывных работ располагался в одном из карьеров Белгородской области, Россия. Съёмка проводилась в марте 2021 года в солнечную погоду в дневное время.

Изображения разрешением 3936x5248 пикселей были получены с использованием беспилотного летательного аппарата DJI Phantom 4 PRO V2.0. Исследования проводились на видеокарте RTX 2080 с объемом видеопамати 12 GB. Размер изображения превышает объем памяти видеокарты, что исключает возможность его обучения в исходном виде. Сеть FPN и предобученные модели большого размера EfficientNet B7 или SE ResNet 101, использованные для сравнения, способны функционировать с изображениями размером 224x224 пикселя на данном объеме видеопамати [10]. Поэтому для обучения было решено разрезать 67 исходных изображений 3936x5248 пикселя на 26197 небольших сегментов 224x224 пикселя. Для каждого сегмента было создано изображение с разметкой с применением One Hot Encoding для работы с НС, которая представлена следующими классами: №1 – поверхность куска руды, №2 – граница куска руды, №3 – фон.

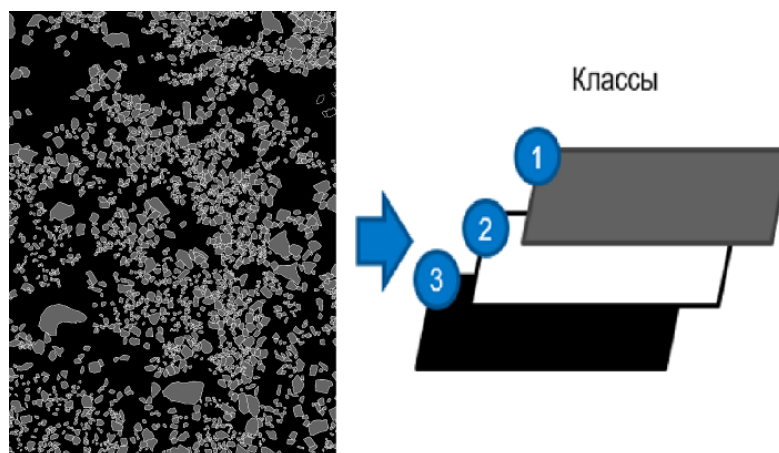


Рис. 2. – Распределение классов на метке

4. Метрики оценки

Для оценки качества сегментации используется метрика IoU (Intersection over Union), одна из самых популярных для семантической сегментации. Она применяется для бинарной и многоклассовой сегментации

и вычисляет долю пересечения между предсказанной и истинной масками к их объединению. Значение IoU варьируется от 0% (полное несовпадение) до 100% (полное совпадение) (формула 1):

$$IoU = \frac{TP}{TP + FP + FN} \quad (1)$$

где:

- **TP** (True Positive) — пересечение истинной и предсказанной масок,
- **FP** (False Positive) — пиксели предсказанной маски, не совпадающие с истиной,
- **FN** (False Negative) — пиксели истинной маски, не совпадающие с предсказанием.

Для многоклассовой сегментации IoU усредняется по всем классам:

$$IoU = \frac{\sum_{i=1}^N IoU_i}{N} \quad (2)$$

где IoU_i — точность для каждого класса, N — количество классов.

5. Результаты и обсуждение

Модели FPN с разными предобученным основаниям семейства EfficientNet была обучены до 10 эпох и протестированы. Сравнительный анализ результатов приведен на рисунке 3.

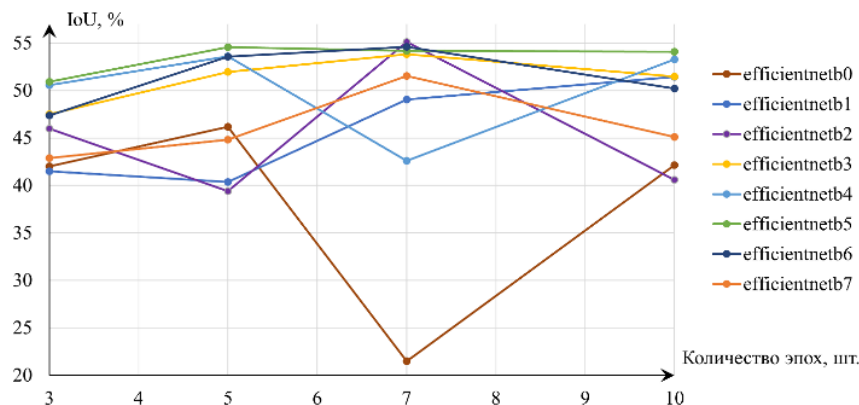


Рис. 3. – График значений IoU по семейству EfficientNet

Максимальное значение показателя IoU, равное 55,1%, соответствует FPN-сети предобученного основания EfficientNet B2 после 7 эпох обучения.

На рисунке 4 представлены сведения о числовом показателе IoU по всему семейству предобученных оснований SE ResNet по каждой эпохе обучения.

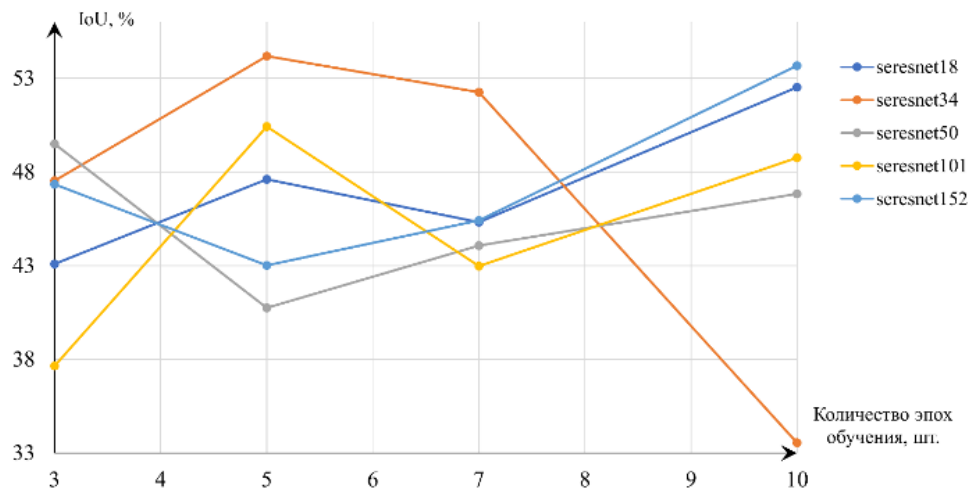


Рис. 4. – График значение IoU по семейству SE ResNet

Максимальное значение показателя IoU, равное 54,2 процента, соответствует FPN-сети на предобученным основанием SEResNet34 после 5 эпох обучения.

Очевидно, что наилучшие результаты достигнуты не на самых мощных моделях предобученных оснований. С одной стороны, результат порядка 55 процентов точности сегментации относительно невысок, с другой он получен на ограниченном объеме данных и видеопамяти графического процессора. Это показывает потенциал применения данных моделей для решения поставленной задачи.

Работа лучших моделей на одном из тестовых изображений приведена на рисунке 5.

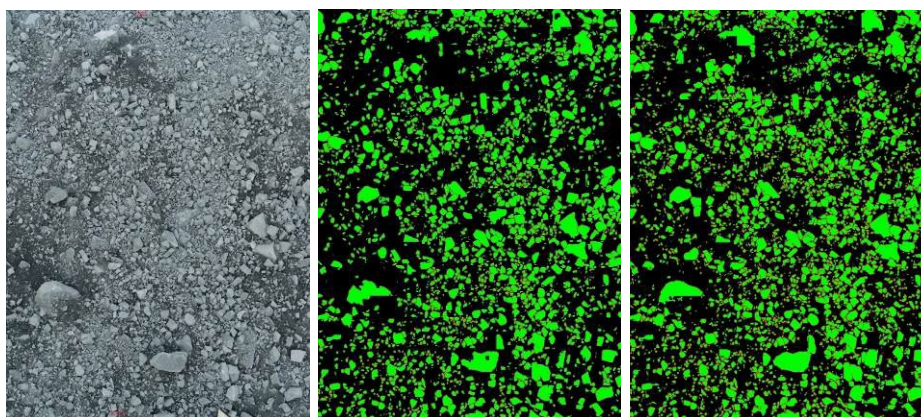


Рис. 5. – Тестовое изображение (слева), результат работы FPN-сети на EfficientNet B2 (7 эпох) (центральное) и на SEResNet34 (справа)

Для определения размеров и подсчета рудных тел предлагается использовать метод оконтуривания объектов из библиотеки OpenCV [10]. Ввиду сложной геометрической формы для нахождения граничных точек объектов может быть использован эллипс. Примеры работы функции оконтуривания с использованием эллипсов показаны на рисунке 6.

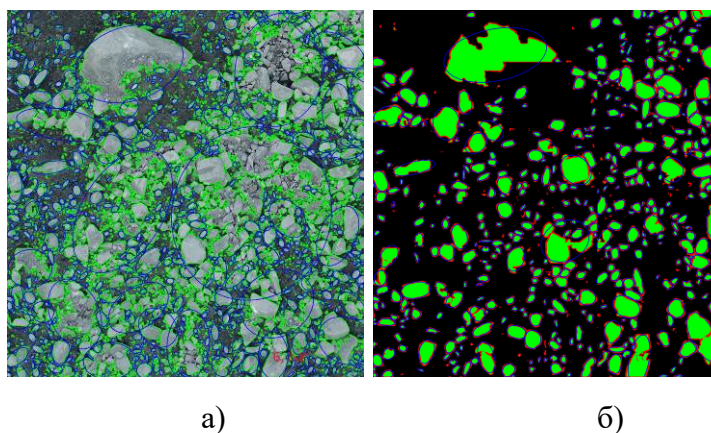


Рис. 6. – Результат работы OpenCV: а) на исходном изображении; б) на сегментированном изображении

Функция оконтуривания дает более точные результаты на сегментированных изображениях. В дальнейшем предполагается

использовать данный подход для определения гранулометрического состава результата буровзрывных работ.

Заключение

В работе полученные результаты демонстрируют потенциальную возможность применения методов глубокого обучения для автоматизации процесса оценки гранулометрического состава руды на основе изображений с БПЛА. Наиболее успешными моделями оказались FPN с предобученными основаниями EfficientNet B2 и SE ResNet 34, достигшие максимальных значений метрики IoU в 55,1% и 54,2% соответственно. Полученные результаты показали перспективу проведения дальнейших исследований для повышения точности сегментации. Это увеличение объема датасета при одновременном увеличении разрешения изображения, подаваемого на нейронную сеть. Эксперименты показали, что сегментированные изображения обеспечивают более точные результаты при применении метода оконтуривания для определения размеров рудных тел.

Литература

1. Буданова Т. Е., Озмидов О. Р., Озмидов И. О. Современные методы изучения гранулометрического состава грунтов //Инженерные изыскания. – 2013. – №. 8. – С. 66.
2. Великанов В.С., Дремин А.В., Чернухин С.А., Ломовцева Н.В. Технологии нейронных сетей в интеллектуальном анализе данных гранулометрического состава взорванных пород. Горная промышленность. 2024;(4):90–94.
3. Wang, W., Li, Q., Zhang, D., Fu, J. Image segmentation of adhesive ores based on MSBA-Unet and convex-hull defect detection. Engineering Applications of Artificial Intelligence. 2023. V. 123. p. 106185.

4. Guo, Q., Yang, S., Wang, Y., Zhang, H. Size measurement of blasted rock fragments based on FRRSNet+. Measurement. 2023. V. 218. p. 113207.
5. Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S. Feature pyramid networks for object detection. Proceedings of the IEEE conference on computer vision and pattern recognition. 2017. pp. 2117-2125.
6. Chang, C. W., Santra, S., Hsieh, J. W., Hendri, P., Lin, C. F. Multi-fusion feature pyramid for real-time hand detection. Multimedia Tools and Applications. 2022. V. 81. No. 9. pp. 11917-11929.
7. Tan M., Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. International conference on machine learning. PMLR, 2019. pp. 6105-6114.
8. He, K., Zhang, X., Ren, S., Sun, J. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. pp. 770-778.
9. Hu J., Shen L., Sun G. Squeeze-and-excitation networks. Proceedings of the IEEE conference on computer vision and pattern recognition. 2018. pp. 7132-7141.
10. Гордиенко О., Полещенко Д. А., Михайлов И. С., Климов В. Е. Определение гранулометрического состава результата буровзрывных работ в карьере с применением нейронных сетей // Инженерный вестник Дона, 2023, № 11. URL: ivdon.ru/ru/magazine/archive/n11y2023/8779.
11. OpenCV - Open Computer Vision Library. – URL: opencv.org (дата обращения 12.09.2024).

References

1. Budanova T. E., Ozmidov O. R., Ozmidov I. O. Inzhenernye izyskaniya. 2013. No. 8. p. 66.
 2. Velikanov V.S., Dremin A.V., Chernuhin S.A., Lomovceva N.V. Gornaya promyshlennost'. 2024. No. 4. pp. 90–94.
-

3. Wang, W., Li, Q., Zhang, D., Fu, J. Engineering Applications of Artificial Intelligence. 2023. V. 123. p. 106185.
4. Guo, Q., Yang, S., Wang, Y., Zhang, H. Measurement. 2023. V. 218. p. 113207.
5. Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S. Proceedings of the IEEE conference on computer vision and pattern recognition. 2017. pp. 2117-2125.
6. Chang, C. W., Santra, S., Hsieh, J. W., Hendri, P., Lin, C. F. Multimedia Tools and Applications. 2022. V. 81. No. 9. pp. 11917-11929.
7. Tan M., Le Q. International conference on machine learning. PMLR, 2019. pp. 6105-6114.
8. He, K., Zhang, X., Ren, S., Sun, J. Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. pp. 770-778.
9. Hu J., Shen L., Sun G. Proceedings of the IEEE conference on computer vision and pattern recognition. 2018. pp. 7132-7141.
10. Gordienko O., Poleshchenko D. A., Mihajlov I. S., Klimov V. E. Inzhenernyj vestnik Dona, 2023, №1. URL: ivdon.ru/ru/magazine/archive/n1y2023/8779.
11. OpenCV - Open Computer Vision Library. URL: opencv.org (accessed: 12/09/24).

Дата поступления: 14.11.2024

Дата публикации: 1.01.2025