

**Метод генерации изображений для повышения качества сегментации
гранулометрических данных**

Усачева Д.М. (Университет ИТМО)

Научный руководитель – кандидат технических наук, **Ефимова В. А.**

(Университет ИТМО)

Введение. Гранулометрический анализ — это процесс определения распределения частиц по размерам в сыпучих материалах, используемый для контроля качества и оптимизации технологических процессов. Автоматизация управления мельницей на основе гранулометрического анализа камней на конвейере повысит эффективность переработки руды, устранив субъективные ошибки ручного регулирования щели дробилки. Применение компьютерного зрения для детекции и сегментации камней позволяет строить их распределение по размерам и автоматически и адаптивно управлять дроблением. Такой подход обеспечивает автоматизированную и точную оценку фракционного состава в реальном времени, оптимизируя раскрытие щели дробилки, снижая энергозатраты и улучшая качество дробления. Для точного анализа гранулометрических данных требуется качественная сегментация изображений. Современные нейросетевые методы — VAE [4], GAN [5] и диффузионные модели [6, 7] — позволяют синтезировать изображения с разметкой, что улучшает обучение моделей сегментации [8]. Основная цель работы — разработка метода генерации изображений для обучения модели сегментации, что повысит точность детекции камней и эффективность управления дроблением руды.

Основная часть. Для решения задачи сегментации гранулометрических данных был проведен сравнительный анализ современных методов генерации изображений, включая VAE, GAN и диффузионные модели. На основе анализа выбран наиболее подходящий метод, который дообучается на ограниченном наборе снимков камней. С его помощью сгенерирован новый набор изображений в количестве не менее 1500 примеров. Для разметки использован метод слабой разметки [1], включающий алгоритм водораздела и модель Segment Anything Model (SAM) [3] в комбинации с метками детекции YOLO [2]. Оценено качество разметки, а также создана небольшая тестовая выборка (не менее 20 изображений) с ручной разметкой для финальной оценки качества. Далее генеративная модель была дообучена с учетом новой модальности (масок сегментации) для генерации пар изображение-маска, что позволило получить не менее 1500 пар данных. Эти данные использованы для обучения сегментационной модели, качество которой оценивалось на вручную размеченном тестовом наборе по метрикам IoU, Dice Score, Precision и Recall. Проведен сравнительный анализ моделей, обученных на слабой разметке и на синтетическом наборе данных, с целью оценки степени улучшения качества сегментации.

Выводы. Разработан метод генерации изображений для сегментации гранулометрических данных, выбрана и дообучена оптимальная генеративная модель. Сгенерировано не менее 1500 пар изображение-маска и обучена модель сегментации.

Проведенный сравнительный анализ показал, что использование сгенерированных данных позволило улучшить качество сегментации: метрика IoU повысилась на 0,2, а F1-Score — на 0,1 по сравнению с моделью, обученной только на слабой разметке.

Список использованных источников:

- [1] Усачева Д.М., Прокопов Е.М., Румянцева М.Ю., Ефимова В.А. Weak Segmentation and Unsupervised Evaluation: Application to Froth Flotation Images//International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, 2025, Vol. 3, pp. 500-507
- [2] YOLO by Ultralytics (Version 8.0.0) <https://github.com/ultralytics/ultralytics>
- [3] Segment Anything. <https://arxiv.org/abs/2304.02643>
- [4] Kingma, D. P., & Welling, M. (2013). Auto-Encoding Variational Bayes.. <https://arxiv.org/abs/1312.6114>
- [5] Goodfellow, I., et al. (2014). Generative Adversarial Networks. *arXiv preprint*. <https://arxiv.org/abs/1406.2661>
- [6] Ho, J., Jain, A., & Abbeel, P. (2020). Denoising Diffusion Probabilistic Models. *arXiv preprint*. <https://arxiv.org/abs/2006.11239>
- [7] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-Resolution Image Synthesis with Latent Diffusion Models. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. <https://arxiv.org/abs/2112.10752>
- [8] Haque, A., & Tajbakhsh, N. (2023). SegGen: Supercharging Segmentation Models with Text2Mask and Mask2Img Synthesis. *arXiv preprint*. <https://arxiv.org/abs/2311.03355>

Усачева Д.М. (автор)

Подпись

Ефимова В.А. (научный руководитель)

Подпись