

Применение методов самообучения для задач сегментации однородных объектов**Крюков А.Д.** (Университет ИТМО), **Румянцева М. Ю.**, (Университет ИТМО)**Научный руководитель – к.т.н., Ефимова В. А.**

(Университет ИТМО)

Введение. Анализ пенной флотации является важнейшей задачей при переработке полезных ископаемых в металлургии. Использование таких подходов к самообучению, как MoCo, SparK, Jigsaw Puzzle и Barlow Twins, помогает нам избегать необходимости в больших наборах размеченных данных. В исследовании подчеркивается уникальный вклад каждого метода с самоконтролем, демонстрируется их влияние на результаты работы моделей. Лучшие из предложенных методов превосходят базовый показатель для помеченного набора данных на 3%. Полученные результаты дают ценную информацию о сильных сторонах и ограничениях использования самоконтролируемого обучения при анализе изображений пены, прокладывая путь для дальнейших достижений в этой области.

Основная часть. Специфика данных пенной флотации состоит в том, что перед нами стоит задача сегментации однородных объектов, а это предполагает большой дисбаланс классов и проблемы с сегментацией на границах объектов. Стоит отметить, что существует довольно много различных алгоритмов обучения без учителя, которые частично решают эту проблему, как например алгоритм водораздела и алгоритм Отсу, которые могут в некоторых случаях выделять объекты. Сегментация при помощи нейросетей требует для большой точности массивные наборы размеченных данных, но для разметки необходимо потратить много времени, поэтому оптимальным является поиск решения, при котором мы можем без потерь в точности использовать меньшие наборы размеченных данных, но с обучением части финальной нейросети на отличной от основной задаче. В качестве решения этой проблемы были предложены методы самообучения, который предполагают обучение на изначально неразмеченных данных. Одной из задач в таком случае является применение к одинаковым объектам разных видов аугментации и последующее обучение нейросети на задачи минимизации разницы в репрезентации одинаковых изначально объектов и максимизации между различными. Одним из самых популярных вариантов является создание мозаик на основе неразмеченных данных и их решение, а также фреймворки основанные на аугментациях, как например используемый нами Barlow Twins, который довольно часто используется при решении задач сегментации [1]. В нашей работе мы исследуем применение данного фреймворка при работе с однородными данными с сильным дисбалансом классов. Еще одним способом самообучения является метод контрастивного обучения, основанный на идее поиска сходств и различий разных объектов, например, фреймворк MoCo[2], в котором мы делаем словарь свойств различных объектов и задачей является минимизацией расстояния между объектом и его свойствами. Последним методом, который мы рассматривали была работа с автоэнкодерами, в ходе работы которого мы маскируем часть фото и далее пытаемся предсказать ее. В своей работе для разработки алгоритма мы использовали такие методы самообучения, как MoCo, Barlow Twins[3], SparK и Jigsaw Puzzle[4]. В качестве основы мы взяли архитектуру Unet [5], так как она часто применяется в решении задач сегментации, и RCA-IUnet [6]. В начале работы мы провели эксперименты по поиску оптимальной loss-функции для обучения нашей модели, по итогам экспериментов наилучшие результаты показала функция бинарной кросс-энтропии. Далее при помощи ранее упомянутых методов был обучен энкодер для Unet и RCA-IUnet при помощи самообучения, далее добавлен декодер и дообучен на размеченном наборе данных,

Выводы. В данной работе мы изучили четыре различных подхода с самоконтролем для улучшения качества сегментации и выбрали два лучших - Barlow Twins и MoCo, причем Barlow Twins показали наилучшие результаты как с точки зрения показателей, так и

визуально (0,63 IoU против 0,54 без него). Хотя этот метод имеет свои ограничения, он значительно улучшает сегментацию по сравнению с базовыми методами. Результаты этого исследования могут быть доработаны и применены не только к данным пенной флотации, но и к изображениям, на которых представлены однородные данные.

Список использованных источников:

- [1] BTSeg: Barlow Twins Regularization for Domain Adaptation in Semantic Segmentation
- [2] Momentum Contrast for Unsupervised Visual Representation Learning
- [3] Barlow Twins: Self-Supervised Learning via Redundancy Reduction
- [4] Unsupervised Learning of Visual Representations by Solving Jigsaw Puzzles
- [5] U-Net: Convolutional Networks for Biomedical Image Segmentation
<https://doi.org/10.48550/arXiv.1505.04597>
- [6] RCA-IUnet: A residual cross-spatial attention guided inception U-Net model for tumor segmentation in breast ultrasound imaging

Крюков А.Д. (автор)

Подпись

Ефимова В.А. (научный руководитель)

Подпись