

УДК 004.93

Кульминский Д. П., Ефимов В. В.

**Semi-supervised сегментация
с использованием нейросетей
в задаче оценки кондиционности керна**

Рекомендовано к публикации доцентом Гришкиным В. М.

1. Введение. Керна – образец горной породы цилиндрической формы. По такому образцу можно получить информацию о геологическом строении недр, вещественном составе горных пород, наличии или отсутствии в них углеводородов. Однако, прежде чем транспортировать образец в кернохранилище или лабораторию, необходимо точно оценить кондиционность керна, т. е. его пригодность для взятия пробы и дальнейшего лабораторного анализа. Некондиционными считаются участки с обилием трещин техногенного (получаются при извлечении из скважины или транспортировке) или естественного происхождения (свойство самой породы) или наличием крупных разломов. Эти участки не участвуют в последующих исследованиях. Соответственно, задача заключалась в разработке инструмента, выполняющего семантическую бинарную сегментацию кондиционности горной породы, т. е. присваивающего каждому пикселю принадлежность к одному из классов: «кондиционный» или «некондиционный».

Традиционно в задаче семантической сегментации в качестве обучающей выборки используются пары из оригинальных изображений и масок. Последние представляют собой изображение, совпадающее с оригиналом по размерности, каждый пиксель которого имеет метку одного из классов, т. е. имеет место переход из какого-либо цветового пространства во множество меток классов. Обычно маски являются результатом ручной разметки специалистов предметной области. Объем выборок из таких пар варьируется от нескольких сотен до

Кульминский Данил Петрович – студент, Санкт-Петербургский государственный университет; e-mail: st064144@student.spbu.ru, тел.: +7(922)194-01-88

Ефимов Владислав Владимирович – студент, Санкт-Петербургский государственный университет; e-mail: st073936@student.spbu.ru, тел.: +7(911)112-58-89

нескольких тысяч примеров. Ввиду того, что предоставить несколько тысяч масок с достаточно чувствительной разметкой кондиционных и некондиционных сегментов за разумное время не представлялось возможным, вынужденной мерой стало использование вместо детальной попиксельной разметки более грубой площадной, представляющей из себя обобщенную оценку качества на том или ином крупном сегменте породы.

2. Метод сегментации изображений керна. Предложенный подход состоял из следующих этапов (см. рис. 1):

1. По исходному изображению составлялась маска классов пород с использованием обученного классификатора, применяемого на окнах, и SS-CAM (Smooth Score Class Activation Map) обученного отдельного классификатора кондиционности [1].
2. Маска SS-CAM бинаризовалась применением пороговой функции попиксельно в зависимости от классов пород.
3. На полученных «грубых» бинарных масках сегментации обучалась модель семантической сегментации U-Net, позволяющая получить финальный результат сегментации в один этап.

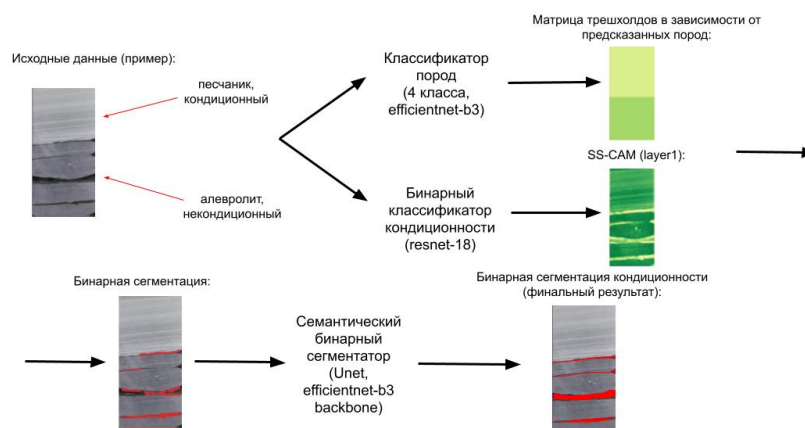


Рис. 1. Схема этапов обучения алгоритма

2.1. Классификатор пород. Задачей классификатора пород было определять следующие классы: «алевролит», «песчаник», «переслаивание пород», «аргиллит». В качестве основы была использована предобученная нейросетевая модель efficientnet-b3 [2].

Предобработка данных для обучения состояла из следующих этапов:

1. Пропорциональное приведение исходного изображения к фиксированной ширине в 256 пикселей.
2. Выбор в качестве сэмплов прямоугольных областей изображения высотой в 512 пикселей и шириной в 256, 50% площади которых принадлежат одному классу породы.
3. Замена черным пикселей, не принадлежащих к целевому классу, т. е. тому, который занимает 50% и более от всей площади сэмпла.

На полученном множестве сэмплов сеть архитектуры efficientnet-b3 дообучалась с использованием алгоритма оптимизации AdamW [3] и кросс-энтропии в качестве функции потерь.

В силу специфики предметной области переиспользование моделей на изображениях иного масштаба крайне нежелательно. Причиной тому является значимость размера вкраплений в качестве визуального признака при определении породы литологических образцов.

В связи с этим в процессе использования модели исходные изображения аналогично предобработке на этапе обучения приводились к ширине в 256 пикселей. Далее изображения разбивались по вертикали на окна высотой в 512 пикселей с пересечением в 60%. Полученные окна использовались в качестве входных данных для модели, на пересечениях вектор вероятностей классов усреднялся.

2.2. SS-CAM бинарного классификатора кондиционности. Поскольку точность разметки не позволяла явно обучить на ней сегментатор кондиционности, для «грубой» сегментации использовалась SS-CAM первого сверточного слоя классификатора кондиционности. Она показывала значимость каждого пикселя в задаче определения класса «кондиционный» или «некондиционный».

Предобработка изображений на этапе обучения была аналогична описанной для классификатора пород. В качестве классификатора использовалась обученная с нуля архитектура resnet-18 [4] с

использованием оптимизатора AdamW и бинарной кросс-энтропией в качестве функции потерь.

На этапе использования SS-CAM классификатора исходные изображения так же приводились к ширине в 256 пикселей. Далее изображения разбивались на окна высотой в 512 пикселей и шириной в 256 с наложением в 60% как по горизонтали, так и по вертикали. Для устранения артефактов крайние 20 пикселей SS-CAM не использовались, для остальных пикселей на пересечениях окон значения соседних SS-CAM усреднялись. Для получения бинарной маски сегментации использовался установленный в зависимости от класса породы порог. На полученной бинарной маске удалялись все сегменты с площадью менее 400 пикселей.

2.3. Нейросеть семантической сегментации. Описанный выше подход является весьма требовательным в плане вычислительных ресурсов – SS-CAM подразумевает итерации по всем фильтрам заданного сверточного слоя для построения оценок важности пикселя исходного изображения для заданного класса. В связи с этим было решено использовать результат работы двух классификаторов для формирования новой обучающей выборки для сегментирующей сети архитектуры U-Net [5], перейдя таким образом к известной задаче бинарной семантической сегментации. Используя пары оригинальных изображений и бинаризованных масок SS-CAM, алгоритм оптимизации AdamW и взвешенную кросс-энтропию в качестве функции потерь (большой штраф за ошибки на пикселях, соответствующих повреждениям, а не куску цельной породы), удалось обучить U-Net распознаванию некондиционных областей на фотографии (метрика сходства предсказаний сегментирующей модели и масок SS-CAM с бинарного классификатора кондиционности, Intersection over union, достигла значения 0,7).

3. Анализ результатов. Предложенный подход позволил добиться решения высокой точности в задаче сегментации без соответствующей разметки. Результат был достигнут благодаря использованию всего объема доступных данных, в том числе разметки пород и информации от экспертов предметной области, позволившей подобрать пороговые значения для алгоритма на первом этапе «грубой» сегментации.

Использование U-Net для построения финального решения существенно повысило скорость обработки изображений.

4. Заключение. В работе продемонстрирован способ решения задачи семантической сегментации в условиях отсутствия детальной попиксельной разметки. Используя технику извлечения карт активаций по классам со сверточных слоев сетей, обученных на классификацию крупных сегментов, возможно либо решить задачу напрямую (после некоторой постобработки), либо на их основе обучить одну из распространенных сегментирующих архитектур.

Литература

1. Haofan Wang, Rakshit Naidu, Joy Michael, Soumya Snigdha Kundu SS-CAM: Smoothed Score-CAM for Sharper Visual Feature Localization // arXiv preprint arXiv:2006.14255. 2020 [Электронный ресурс]: URL:<https://arxiv.org/abs/2006.14255> (дата обращения: 10.10.2020).
2. Mingxing Tan, Quoc V. Le EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks // arXiv preprint arXiv:1905.11946. 2020 [Электронный ресурс]: URL:<https://arxiv.org/abs/1905.11946> (дата обращения: 15.10.2020).
3. Loshchilov I., Hutter F. Decoupled Weight Decay Regularization // arXiv preprint arXiv:1711.05101. 2019 [Электронный ресурс]: URL:<https://arxiv.org/abs/1711.05101> (дата обращения: 14.10.2020).
4. Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun Deep Residual Learning for Image Recognition // arXiv preprint arXiv:1512.03385. 2015 [Электронный ресурс]: URL:<https://arxiv.org/abs/1512.03385> (дата обращения: 05.10.2020).
5. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation // arXiv preprint arXiv:1505.04597. 2015 [Электронный ресурс]: URL:<https://arxiv.org/abs/1505.04597> (дата обращения: 14.09.2020).