

ИССЛЕДОВАНИЕ МЕТОДОВ АУГМЕНТАЦИИ В ЗАДАЧЕ СЕГМЕНТАЦИИ КАМНЕЙ НА КОНВЕЙЕРЕ ПРЕДПРИЯТИЯ ГОРНОЙ ПРОМЫШЛЕННОСТИ

STUDY OF AUGMENTATION METHODS IN THE PROBLEM OF STONE SEGMENTATION ON A CONVEYOR BELT OF A MINING ENTERPRISE

V. Kalashnikov

Summary. This paper examines various image augmentation methods in the problem of segmenting stones on a conveyor belt of a mining industry. The use of the deep learning algorithm MaskRCNN and the PyTorch library to solve the problem is considered. Using the IoU metric as an example, the quality of augmentation under various transformations is analyzed. The most and least effective approaches were identified and their qualitative analysis was carried out.

Keywords: mining, quality control, artificial intelligence, neural networks, deep learning, augmentation.

Калашников Владимир Андреевич
Аспирант, Финансовый Университет
при Правительстве РФ (Москва)
vakalashnikov@fa.ru

Аннотация. В данной работе исследуются различные методы аугментации изображений в задаче сегментации камней на конвейере предприятия горной промышленности. Рассмотрено применение алгоритма глубокого обучения MaskRCNN и библиотеки PyTorch для решения поставленной задачи. На примере метрики IoU проанализировано качество аугментации при различных преобразованиях. Выявлены наиболее и наименее эффективные подходы и проведен их качественный анализ.

Ключевые слова: горная промышленность, контроль качества, искусственный интеллект, нейронные сети, глубокое обучение, аугментация.

Введение

Последнее десятилетие ознаменовано активным внедрением методов искусственного интеллекта (ИИ) в различные сферы общественной жизни. По данным АНО «Цифровая экономика» [1] за период 2020–2022 года прирост финансирования проектов в сфере ИИ составил 170 %. При этом более 75 % компаний уже интегрировали использование глубокого и машинного обучения в свои бизнес-стратегии. Активнее всего внедряют передовые технологии в сферах ритейла, беспилотных автомобилей и потребительских товаров и услуг. Таким образом, в ближайшее десятилетие будет проведена цифровая трансформация многих отраслей экономики, в том числе и промышленного производства.

Автоматизация затронет и такую важную отрасль российской экономики, как горная промышленность. На данный момент уже существуют отечественные компании — разработчики ИИ, активно внедряющие свои решения в этой области, например ГК Цифра.

Безусловно, прогресс в цифровизации производства связан не только с работой коммерческих компаний в данной отрасли, но и с проведением научных и методологических исследований. Исходя из всего сказанного выше можно заключить, что тема данной статьи — ана-

лиз методов аугментации данных для сегментации продукции горного предприятия является актуальной.

Аугментация данных

В общем случае, аугментацией данных называют пополнение набора данных для обучения искусственными примерами, сгенерированными на базе уже имеющихся с помощью различных методов. В задачах компьютерного зрения процесс аугментации состоит в использовании различных преобразований, например поворотов изображения, применения различных фильтров и пр. В использованной авторами библиотеке PyTorch имеется достаточно большой набор встроенных преобразований [6], позволяющий исследователям достаточно быстро проверить эффективность тех или иных видов аугментации.

Общая идея данного метода состоит в следующем, если начальный набор данных позволяет извлечь из изображений некоторые паттерны, то его расширение может модифицировать имеющиеся паттерны таким образом, что обучение нейронной сети улучшится.

Приведем конкретный пример: пусть наш набор данных состоит из изображений камней, и задача состоит в их сегментации. Очевидно, начальное положение камней на изображении не несет никакой особой информа-

ционной нагрузки с точки зрения рассматриваемой задачи, т.е. если бы мы сфотографировали камни с другого ракурса или под другим углом, то постановка задачи бы не изменилась, однако границы камней на изображении были бы изменены под действием некоторого проективного преобразования, что изменило бы процесс обучения нейронной сети. Применяя к имеющимся изображениям цифровые преобразования изображений, мы можем добиться существенного изменения видимой формы камней, что позволит нам получить гораздо более широкий набор паттернов, задействованный в обучении. Эту роль и выполняет аугментация изображений.

Анализ литературы (ист. 18, ист. 7, ист. 15, моя)

Задача сегментации породы различных видов затрагивалась отраслевыми исследователями в области машинного обучения в последние годы. В частности, в статье [2] исследователи применили архитектуру HCN для решения задачи сегментации кусков цемента на изображении. Примечательно, что авторы использовали для обучения нейронной сети небольшой набор данных, состоящий из 150 изображений, однако применение методов аугментации (дополнения) данных с помощью различных преобразований позволило достичь довольно высокой точности сегментации: доля корректно определенных пикселей составила более 90 %. Тем не менее, авторы приводят лишь свои итоговые результаты, не останавливаясь подробно на изменении метрик в зависимости от применяемых преобразований.

Каримпоули и др. [3] решали задачу определения вида и характеристик камней на изображении и добились доли верных ответов более 96 % в задаче многоклассовой сегментации. Примечательно, что для обучения нейронной сети (использована архитектура SegNet) был использован набор данных, состоящий всего из 20 изображений. Используя методы аугментации данных, исследователи сгенерировали датасет из 22000 изображений, распределенных в отношении 16000/4000/2000 для обучения, тестирования и валидации соответственно. Данный пример демонстрирует наличие серьезного потенциала в применении исследуемого метода к данному классу задач.

Примером успешного применения данного подхода может служить работа Лианга и др. [4]. В ней исследователи использовали одну из самых современных и эффективных архитектур для сегментации — U-Net. Для ее обучения были использованы четыре изображения камней разрешением 3648x2432. Авторы создали маски для каждого из изображений, разметив их вручную и разрезав на отдельные изображения. Затем были использованы некоторые методы аугментации (дополнения) данных: повороты, отражения, зашумление и изменения яркости. В результате обучения авторы получают точность классификации пикселей более 98 %.

Как можно видеть из приведенного обзора, методы аугментации данных позволяют получать устойчивые результаты в задачах сегментации породы даже с использованием небольших (менее 100) изображений. Однако ни в одном из исследований не было приведено качественного анализа влияния применяемых преобразований на итоговые результаты. Исследование данного вопроса и будет являться целью данной работы.

Исследование

В качестве исходного набора данных был использован набор из 100 фотографий щебня на конвейере, снятых в разрешении 1920*1080 на горнодобывающем предприятии. Каждая из фотографий представляет собой изображение движущегося сырья для производства щебня, то есть набора отдельных камней. Расстояние от матрицы камеры до ленты конвейера составляло примерно 130 см, а скорость движения породы — 1,5 м/с.

Пример изображения представлен на Рис. 1.



Обучение проводилось с помощью использования алгоритма сверточных нейронных сетей — MaskRCNN с использованием аргументированных данных к описанному датасету. Более подробное изложение применяемых методов можно найти в статье авторов [5].

В результате проведения исследования набора преобразований были получены следующие результаты:

Таблица 1.

Сравнение различных преобразований при аугментации датасета из 100 изображений

Преобразование	Метрика IoU
Gaussian blur	0.867
GrayScale	0.902
Random Horizontal Flip	0.904
Sharpness(10)	0.884
Solarize(0.5)	0.896
RandomCrop(size=(128, 128))	0.902
RandomCrop(size=(256, 256))	0.907
permutate_channels(permutation=[2, 1, 0])	0.706

Отметим эффективность преобразования $RandomCrop(size=(256, 256))$. Принцип действия данного преобразования заключается в вырезании из изображения случайным образом квадрата со стороной 256 пикселей, на котором в результате будет обучаться нейронная сеть.

Точно определить, почему то или иное преобразование является наиболее эффективным затруднительно, поскольку оно затрагивает внутренние процессы нейронной сети и много миллиардов арифметических вычислений, однако можно предположить, что, в данном случае эффективность метода обусловлена большой вариацией камней на изображении. Вырезание случайных квадратов из картинки позволяет сверточным слоям эффективно научиться распознавать камни в различных масштабах, что, в свою очередь приводит к росту метрик качества алгоритма.

Худшую эффективность демонстрируют следующие два преобразования:

Gaussian blur — гауссово размытие и `permutate_channels` — перестановка каналов RGB изображения. Можно выдвинуть гипотезу — низкая эффективность данных методов обусловлена большой важностью цветовых характеристик породы и четкой различимости границ.

Таким образом, в результате исследования было выяснено, что в задаче сегментации камней на изображении ленты конвейера эффективным является применение методов изменения масштаба и случайных поворотов изображения.

ЛИТЕРАТУРА

1. Белая книга цифровой экономики 2022 [Электронный ресурс] // Автономная некоммерческая ассоциация "Цифровая экономика" URL: http://academim.org/art/pan1_2.html (дата обращения: 18.11.2023).
2. Chen H., Yi J., Guiqiang L., Biao C. Automated cement fragment image segmentation and distribution estimation via a holistically-nested convolutional network and morphological analysis. // *Powder Technology*. 2018. DOI: 339. 10.1016/j.powtec.2018.08.015.
3. Karimpouli S., Tahmasebi P. Segmentation of digital rock images using deep convolutional autoencoder networks, // *Computers & Geosciences*. Vol. 126. 2019, Pp. 142–150. ISSN 0098-3004. DOI: 10.1016/j.cageo.2019.02.003.
4. Liang Z., Zhihong N., Aijun A., Wang X. A particle shape extraction and evaluation method using a deep convolutional neural network and digital image processing. // *Powder Technology*. Vol. 353. 2019. DOI: 10.1016/j.powtec.2019.05.025.
5. Калашников, В.А. Приложения компьютерного зрения в горнодобывающей промышленности / В.А. Калашников, В.И. Соловьев // *Прикладная информатика*. — 2023. — Т. 18, № 1(103). — С. 4–21. — DOI 10.37791/2687-0649-2023-18-1-4-21. — EDN RUJCOY.
6. PyTorch transforms [Электронный ресурс] // PyTorch documentation URL: <https://pytorch.org/vision/main/transforms.html> (дата обращения: 18.11.2023).

© Калашников Владимир Андреевич (vakalashnikov@fa.ru)

Журнал «Современная наука: актуальные проблемы теории и практики»