

Сегментация неоднородностей на продуктах и минералах

Тучин В.С. (ИТМО),

Научный руководитель – кандидат технических наук, ассистент ФИТиП Ефимова В.А.
(ИТМО)

Введение. SWIR-камеры позволяют находить неоднородности на различных целевых объектах - грибковые инфекции и дефекты в продуктах, повреждения и количество содержимого в упаковках. Также такие устройства способны различать минеральные слои в зернах. Однако такие камеры имеют высокую стоимость ввиду дороговизны их комплектующих - матриц и фильтров. Устройство со сниженной стоимостью, но с тем же SWIR-функционалом, можно создать с помощью системы визуализации на основе метода однопиксельной камеры (SPC), однако для обработки кадров низкого разрешения необходимо применять методы глубокого обучения, в том числе модели сегментации для нахождения неоднородностей в исследуемых объектах.

Основная часть. Для упрощения процесса разметки датасета из SWIR изображений была рассмотрена модель Segment Anything 2.1 (SAM-2.1) [1]. Изображения SWIR камеры всегда получаются в серой шкале яркостей. Дообучение происходило на 1394 изображениях такого же формата. Датасет был составлен из 9 различных наборов изображений продуктов с дефектами, в том числе использовался MVТес AD [2]. Дообученная модель научилась сегментировать дефекты на тестовых изображениях продуктов. Базовая же версия определяла только сам объект и фон. Для нахождения различных слоев минералов были отобраны 97 изображения с сегментационными масками из датасета MinerallImage5k [3], однако базовая и дообученная модель SAM 2.1 не смогла решить поставленную задачу на тестовых изображениях. Минеральные вкрапления имеют более сложную морфологическую структуру, чем дефекты продуктов, что накладывает дополнительные условия к размеру и разнообразию тренировочной выборки.

В качестве целевых моделей сегментации были рассмотрены DeepLabV3-ResNet50 и DeepLabV3-ResNet101. Дообучение на датасете PlantSeg [4], содержащим 115 классов дефектов, в ч/б формате для имитации SWIR кадров дало рост целевых метрик в сравнении с теми, какие получены в оригинальной статье, – mIoU с 0.17 до 0.22 и mAcc с 0.38 до 0.8 для DeepLabV3-ResNet50. Для собранного датасета из 1394 изображений mIoU и mAcc равны 0.37 и 0.9 соответственно.

Для упрощения нахождения неоднородностей моделями DeepLabV3-ResNet50 и DeepLabV3-ResNet101 все классы дефектов в собранном датасете и PlantSeg были сведены к одному классу “defect”. Такое преобразование улучшило работу моделей по нахождению сегментационных масок дефектов, что подтверждается высокими значениями целевых метрик mIoU, mAcc, Dice score и mAP.

Выводы. Проведена успешная сегментация с подсчетом метрик на изображениях продуктов с различными дефектами, имитирующих SWIR кадры. Полученные результаты далее помогут в работе с реальными SWIR изображениями целевых объектов.

Список использованных источников:

1. Ravi N. et al. Sam 2: Segment anything in images and videos //arXiv preprint arXiv:2408.00714. – 2024.
2. Bergmann P. et al. MVТес AD--A comprehensive real-world dataset for unsupervised anomaly detection //Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. – 2019. – С. 9592-9600.

3. Nesteruk S. et al. MineralImage5k: A benchmark for zero-shot raw mineral visual recognition and description //Computers & Geosciences. – 2023. – T. 178. – C. 105414.
4. Wei T. et al. Plantseg: A large-scale in-the-wild dataset for plant disease segmentation //arXiv preprint arXiv:2409.04038. – 2024.