

Секция «Искусственный интеллект и анализ данных в космических исследованиях»

Исследование эффективности синтетических обучающих данных для интеллектуальной системы технического зрения робота-манипулятора**Научный руководитель – Карташев Всеволод Владимирович****Саакян Михаил Николаевич***Студент (специалист)*Московский государственный университет имени М.В.Ломоносова, Факультет
космических исследований, Москва, Россия*E-mail: saakyan.mih@gmail.com*

С развитием робототехники и автоматизации промышленные роботы-манипуляторы становятся незаменимым инструментом для выполнения широкого спектра производственных задач. Их использование существенно повышает эффективность, точность и скорость операций, а также обеспечивает возможность работы в условиях, недоступных для человека. Одним из современных методов реализации системы технического зрения робота-манипулятора является применение нейронных сетей.

Однако, для успешного обучения нейронных сетей требуются качественные обучающие данные. Традиционные методы сбора данных зачастую оказываются чрезвычайно трудоёмкими и подвержены ошибкам, связанным с человеческими неточностями при разметке. Для преодоления этих ограничений предлагается использовать синтетические обучающие данные, сгенерированные с помощью средств 3D-моделирования. Такой подход позволяет не только существенно сократить временные затраты на сбор и разметку датасета, но и повысить точность создаваемых обучающих выборок.

Кроме того, синтетические обучающие данные открывают возможность создания датасетов для тех задач, для которых сбор реальных данных затруднён природой самой задачи, как, например, в медицине или космической промышленности. Это позволяет исследователям и инженерам разрабатывать и тестировать модели в областях, где получение экспериментов с реальными данными может быть дорогостоящим или даже невозможным.

В рамках данной работы исследуется эффективность использования синтетических обучающих данных на примере задачи распознавания картонных коробок. Для генерации обучающих выборок были созданы два датасета посредством среды 3D-моделирования Blender — Synth-Box-v1 и Synth-Box-v2. Synth-Box-v2 отличается от первой версии увеличенным разнообразием видов картонных коробок, а также разнообразием задних фонов. Каждый из датасетов включает 1500 обучающих примеров и 500 примеров для валидации. Для оценки качества полученных моделей применялся реальный датасет Box-IS, состоящий из 543 примеров фотографий картонных коробок [1].

В эксперименте использовались две модели instance segmentation — YOLOv8m-seg и YOLOv11m-seg [2]. Обе модели обучались как с начальными весами, так и с предобученными на датасете COCO [3], что позволило оценить влияние предварительного обучения на качество конечного решения. Результаты вычислительных экспериментов приведены в Таблице 1.

Таблица 1: Метрики качества для моделей YOLOv8 и YOLOv11 на датасете Box-IS

Model	amp; Dataset	amp; Box P	amp; Box R	amp; Box mAP50	amp; Box mAP50-95	amp; Mask P	amp; Mask R	amp; Mask mAP50	amp; Mask mAP50-95
YOLOv8 (COCO)	amp; Synth-Box-v1	amp; 0.448	amp; 0.275	amp; 0.273	amp; 0.162	amp; 0.439	amp; 0.262	amp; 0.252	amp; 0.137
YOLOv8	amp; Synth-Box-v1	amp; 0.346	amp; 0.252	amp; 0.231	amp; 0.103	amp; 0.333	amp; 0.235	amp; 0.204	amp; 0.0795
YOLOv11 (COCO)	amp; Synth-Box-v1	amp; 0.410	amp; 0.279	amp; 0.277	amp; 0.160	amp; 0.401	amp; 0.265	amp; 0.259	amp; 0.136
YOLOv11	amp; Synth-Box-v1	amp; 0.346	amp; 0.267	amp; 0.227	amp; 0.0986	amp; 0.332	amp; 0.252	amp; 0.202	amp; 0.0783
YOLOv8 (COCO)	amp; Synth-Box-v2	amp; 0.604	amp; 0.397	amp; 0.454	amp; 0.329	amp; 0.604	amp; 0.393	amp; 0.441	amp; 0.295
YOLOv8	amp; Synth-Box-v2	amp; 0.481	amp; 0.377	amp; 0.393	amp; 0.223	amp; 0.471	amp; 0.348	amp; 0.353	amp; 0.184
YOLOv11 (COCO)	amp; Synth-Box-v2	amp; 0.589	amp; 0.448	amp; 0.485	amp; 0.327	amp; 0.593	amp; 0.435	amp; 0.463	amp; 0.294
YOLOv11	amp; Synth-Box-v2	amp; 0.428	amp; 0.326	amp; 0.310	amp; 0.162	amp; 0.409	amp; 0.305	amp; 0.274	amp; 0.129

Несмотря на то, что полученные метрики качества ещё оставляют пространство для улучшения, результаты экспериментов демонстрируют, что увеличение разнообразия синтетических данных приводит к улучшению метрик, особенно в случаях, когда модели обучались с начальными весами. В дальнейшем планируется усовершенствовать синтетический датасет, чтобы он точнее имитировал реальные сценарии, с которыми может столкнуться модель. Так же планируется применить данный метод создания обучающих данных для решения задачи распознавания металлических деталей, для которой на данный момент отсутствуют размеченные реальные датасеты.

Список литературы

- [1] Aecheon Jung, Soyun Choi, Junhong Min, and Sungeun Hong, “IAM: Enhancing RGB-D Instance Segmentation with New Benchmarks,” *arXiv preprint arXiv:2501.01685*, 2025. [Online]. Available: <https://arxiv.org/abs/2501.01685>.
- [2] Glenn Jocher, Jing Qiu, and Ayush Chaurasia, *Ultralytics YOLO* (version 8.0.0, January 2023). Licensed under AGPL-3.0. [Online]. Available: <https://github.com/ultralytics/ultralytics>.
- [3] Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, and Piotr Dollár, *Microsoft COCO: Common Objects in Context*, 2015. [Online]. Available: <https://arxiv.org/abs/1405.0312>.