

УДК 004.896

ИССЛЕДОВАНИЕ АЛГОРИТМОВ АУГМЕНТАЦИИ ИЗОБРАЖЕНИЯ ДЛЯ ЗАДАЧ ДЕТЕКТИРОВАНИЯ НА МАЛЫХ НАБОРАХ ДАННЫХ

Поляков А.А. (ИТМО), Ткачѐв И.Ю. (ИТМО), Васильев В.С. (ИТМО)

Научный руководитель – кандидат технических наук, доцент Бжихатлов И.А.
(ИТМО)

Введение. В ряде задач компьютерного зрения сбор и аннотирование больших объемов данных является затруднительным или практически невозможным [1]. Это характерно, например, для детекции брака в производстве, поскольку каждый дефект уникален, или для обнаружения утечек и поломок в трубопроводах, так как такие события редки и непредсказуемы. В подобных ситуациях минимальный возможный объем данных составляет 100-150 изображений. Однако малый объем данных может привести к недостаточной обобщающей способности модели [2]. Одним из решений данной проблемы является аугментация изображений [3], позволяющая искусственно увеличить объем данных и улучшить метрики модели.

Основная часть. В работе исследуется влияние различных подходов к аугментации изображений на качество детекции объектов при малом объеме обучающего набора. В качестве базовой модели выбрана предобученная на датасете COCO нейронная сеть YOLOv8 [4] и YOLOv10 [5].

Изначальный набор изображений взят из другого датасета - OpenImagesV6 [6], который увеличивается как минимум в 10 раз с применением трёх типов аугментаций в разных пропорциях:

- Цветовые искажения (изменение яркости, контраста, насыщенности [7]);
- Геометрические преобразования (вращение, эластичное и оптическое искажения и др.);
- Смешанные аугментации (Мозаика изображений [8], смешивание изображение с помощью прозрачности [9], случайное удаление фрагментов изображения и др.).

Обучение моделей проводилось при одинаковых гиперпараметрах с единственным отличием в использовании различных комбинаций аугментаций. Оценка качества моделей осуществляется на основе метрик точности (mAP, IoU), а также визуального анализа предсказаний.

Выводы. В результате исследования выявлены наиболее эффективные комбинации аугментаций, позволяющие улучшить детекцию объектов на малых наборах данных. Установлено, что оптимальная стратегия включает баланс между цветовыми, геометрическими и композиционными трансформациями.

Список использованных источников:

1. Mingle Xu, Sook Yoon, Alvaro Fuentes, Dong Sun Park A Comprehensive Survey of Image Augmentation Techniques for Deep Learning // arXiv preprint:2205.01491, 2023.
2. Hongyi Zhang Moustapha Cisse, Yann N. Dauphin, David Lopez-Paz mixup: Beyond empirical risk minimization // arXiv preprint:1710.09412, 2018.
3. Enhance dataset using Albuementations [Электронный ресурс]. Режим доступа: <https://docs.ultralytics.com/integrations/albuementations/> (дата обращения 10.02.2025).
4. YOLOv8 Documentation [Электронный ресурс]. Режим доступа: <https://docs.ultralytics.com/models/yolov8/> (дата обращения 24.01.2025).
5. YOLOv10 Documentation [Электронный ресурс]. Режим доступа: <https://docs.ultralytics.com/models/yolov10/> (дата обращения 24.01.2025).

6. Open Images Dataset V6 [Электронный ресурс]. Режим доступа: <https://storage.googleapis.com/openimages/web/download.html> (дата обращения 12.02.2025).
7. Alexey Bochkovskiy, Chien-Yao Wang, Hong-Yuan Mark Liao YOLOv4: Optimal Speed and Accuracy of Object Detection // arXiv preprint arXiv:2004.10934, 2020.
8. Rabin Dulal, Lihong Zheng, Muhammad Ashad Kabir, Shawn McGrath, Jonathan Medway, Dave Swain, Will Swain Automatic Cattle Identification using YOLOv5 and Mosaic Augmentation: A Comparative Analysis // arXiv preprint arXiv:2210.11939, 2022.
9. Zhi Zhang, Tong He, Hang Zhang, Zhongyue Zhang, Junyuan Xie, Mu Li Bag of Freebies for Training Object Detection Neural Networks // arXiv preprint: 1902.04103, 2019.
10. Wei Jiang, Na Ying Improve Object Detection by Data Enhancement based on Generative Adversarial Nets // arXiv preprint: 1903.01716, 2019.

Автор _____ Поляков А.А.

Научный руководитель _____ Бжихатлов И.А.