

АНАЛИЗ АЛГОРИТМОВ АУГМЕНТАЦИИ ИЗОБРАЖЕНИЙ ДЛЯ ОБУЧЕНИЯ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ

Небаба С.Г.

Томский политехнический университет, к.т.н., доцент ОИТ ИШИТР, e-mail: stepanlfx@tpu.ru

Аннотация

В работе проводится сравнительный анализ 3 алгоритмов аугментации изображений (поворот, отражение, мозаичное размещение) для повышения эффективности обучения сверточных нейронных сетей. Получены результаты точности детектирования объектов с помощью сверточной нейронной сети YOLOv7-Tiny, обученной на нескольких аугментированных датасетах, в задаче обнаружения объектов в воздушном пространстве.

Ключевые слова: детектирование объектов в воздушном пространстве, сверточная нейронная сеть, аугментация изображений, YOLO.

Введение

Детектирование летающих объектов (в контексте обнаружения объекта и распознавания его класса) является одной из ключевых задач в сфере контроля воздушного пространства. Быстродействие в решении этой задачи является критичным условием (вплоть до требований работы средств распознавания в режиме реального времени), что ведет к активному использованию систем детектирования (обнаружения) и классификации объектов, основанных на алгоритмах компьютерного зрения и машинного обучения, в том числе на сверточных нейронных сетях (СНС) [1,2].

Точность классификации объектов с помощью СНС во многом зависит от архитектуры сети и датасета, на котором производится ее обучение. В рамках задачи классификации летающих объектов необходим датасет, содержащий изображения различных пилотируемых и беспилотных летательных аппаратов (БПЛА), птиц и других объектов, которые могут быть встречены в воздушном пространстве.

Одна из основных проблем, с которыми можно столкнуться в процессе подготовки датасета, связана с недостаточным количеством изображений объектов каждого класса. Для решения этой задачи используется ряд алгоритмов аугментации, которые показывают свою эффективность в тех или иных условиях обучения СНС.

Основная цель работы состоит в исследовании эффективности различных алгоритмов аугментации датасета, который составлен из небольшого числа изображений объектов в воздушном пространстве.

В качестве СНС для исследования алгоритмов аугментации выбрана YOLOv7-Tiny [3], которая показывает хорошее соотношение точности и скорости классификации объектов в воздушном пространстве [4]. Исследование ведется при детектировании на изображениях объектов 4 классов: птицы, квадрокоптеры, БПЛА самолетного типа и неизвестные объекты.

Описание исходного датасета

В качестве исходного датасета с малым числом изображений использовался набор изображений, на которых представлено 4 класса объектов: bird (птица), uav quadcopter type (БПЛА-квадрокоптеры), uav plane type (БПЛА самолетного типа), undefined (другие, неизвестные объекты). Общий объем датасета - 1946 изображений.

Изображения в датасете сформированы таким образом, чтобы присутствовали как примеры одиночных объектов на изображении, так и примеры множества объектов в кадре. Однако при этом число изображений каждого класса недостаточно для устойчивого обучения нейронной сети. Общее число объектов по классам следующее:

- bird – 387;
- uav quadcopter type – 1141;
- uav plane type – 586;
- undefined – 672.

Датасет разделен на следующие выборки:

- обучающая – 1380 изображений (70 %);

- проверочная (валидационная) – 389 изображений (20 %);
- тестовая – 190 изображений (10 %).

Примеры 4 изображений из датасета представлены на рис. 1.



Рис. 1. Примеры 4 изображений из датасета

Аугментация изображений датасета

Аугментация изображений – это процесс внесения разнообразия в датасет путем применения различных трансформаций к исходным изображениям и добавление трансформированных изображений в датасет. Целью аугментации является улучшение обобщающей способности модели СНС, расширение разнообразия данных для более эффективного обучения [5].

Алгоритмы аугментации данных повышают способность моделей СНС к точному распознаванию объектов с большой степенью изменчивости при малом объеме выборки. Объекты в воздушном пространстве относятся к наиболее динамичным и визуально изменчивым объектам.

К основным алгоритмам аугментации изображений, которые легко автоматизируются, относятся:

- 1) изменение размера (масштабирование);
- 2) поворот;
- 3) отражение (по горизонтали или по вертикали);
- 4) смещение;
- 5) изменение яркости и/или контраста;
- 6) добавление шума;
- 7) изменение оттенка и насыщенности;
- 8) обрезка;
- 9) мозаичное размещение.

Следует иметь в виду, что неправильное или чрезмерное применение аугментации может негативно повлиять на качество распознавания объектов. Так, если применять слишком много различных алгоритмов аугментаций, это может привести к потере изначальной информации на изображениях. Например, повороты, изменения масштаба и другие трансформации могут сделать объекты менее различимыми, что понизит эффективность обучения СНС [5].

В качестве исследуемых алгоритмов аугментации изображений для обучения СНС YOLOv7-Tiny были выбраны алгоритмы поворота изображения, зеркального отражения и мозаичного размещения (пример изображений представлен на рис. 2., как наиболее популярные и универсальные среди всех автоматизируемых алгоритмов аугментации [6].



Рис. 2. а – повернутое изображение, б – мозаичное размещение нескольких изображений

Для автоматизированной аугментации изображений использовались инструменты Roboflow [7]. В результате из исходного набора изображений (OriginDataset) были получены 4 версии аугментированных датасетов, число изображений в которых следующее:

1. Датасет, дополненный зеркально отраженными изображениями (FlipDataset) – 3013.
2. Датасет, дополненный повернутыми (на угол от -15 до 15 градусов) изображениями (RotationDataset) – 4674.
3. Датасет, дополненный мозаичным размещением изображений (MosaicDataset) – 4719.
4. Датасет, дополненный всеми вышеперечисленными методами аугментации (AugmentedDataset) – 4719.

Созданные с использованием аугментации изображений датасеты использовались для обучения СНС YOLOv7-Tiny.

Описание архитектуры СНС YOLOv7-Tiny

YOLOv7 – одна из наиболее актуальных СНС класса YOLO, которые отличаются высокими показателями точности и скорости детектирования объектов на изображениях. СНС YOLOv7 содержит 132 сверточных слоя и отличается от предыдущих версий YOLO рядом оптимизаций (таких как расширенное эффективное объединение слоев, Extended efficient layer aggregation networks, E-ELAN), что позволяет добиться большего быстродействия при той же точности обнаружения и классификации объектов [3].

YOLOv7-Tiny – версия СНС YOLOv7, с сокращенным числом сверточных слоев (87) и рядом других оптимизаций. Малый размер этой сети делает ее перспективной в задаче детектирования объектов в режиме реального времени (при обработке 30 и более кадров в секунду) даже на мобильных системах компьютерного зрения. При этом точность этой СНС не уступает более крупным и тяжеловесным СНС класса YOLO прошлых поколений (например, YOLOv4) в задаче детектирования объектов в воздушном пространстве [4].

Результаты исследования точности детектирования объектов на аугментированных датасетах

Наиболее устоявшейся метрикой для оценки точности классификации объектов на изображениях с помощью СНС является метрика Average Precision (AP) [8]. Она рассчитывается для каждого класса объектов отдельно, следовательно, необходимо рассчитать количество значений AP равное 4 (по числу рассматриваемых в поставленной задаче классов объектов). Значения AP усредняются для получения метрики mAP (mean AP) – среднего значения AP по всем классам. Метрика AP рассчитывается при заданном пороге Intersection over Union, IoU [9]. В ходе исследования использовался единый порог IoU=0.5.

Все результаты по точности и быстродействию исследуемых СНС были получены с помощью фреймворка PyTorch [10].

Число эпох обучения у СНС YOLOv7-Tiny было равным 100, доверительный порог, conf_thres, был выбран равным 0,25.

Результаты тестирования СНС YOLOv7-Tiny по метрикам AP и mAP на тестовой выборке оригинального датасета (OriginDataset) представлены в таблице 1.

Таблица 1

Точность детектирования объектов с помощью СНС YOLOv7-Tiny, обученной на различных датасетах

Датасет	bird, AP _{0.5}	uav plane type, AP _{0.5}	uav quadcopter type, AP _{0.5}	undefined, AP _{0.5}	mAP _{0.5}
OriginDataset	0,543	0,638	0,753	0,697	0,658
FlipDataset	0,575	0,643	0,768	0,749	0,684
RotationDataset	0,579	0,651	0,764	0,722	0,679
MosaicDataset	0,537	0,614	0,743	0,681	0,644
AugmentedDataset	0,541	0,654	0,771	0,722	0,678

Заключение

Сравнительный анализ всех полученных результатов по оценке точности детектирования объектов на изображениях различных датасетов с помощью СНС YOLOv7-Tiny позволил сделать следующие выводы.

Точность детектирования объектов на изображениях тестовой выборки с помощью СНС YOLOv7-Tiny выше в среднем при аугментациях с помощью отражения (FlipDataset) и поворота (RotationDataset). При этом аугментация с помощью мозаичного размещения (MosaicDataset) влияет на итоговую точность распознавания негативно как сама по себе, так и в сочетании с другими методами аугментации (AugmentedDataset). Средняя точность по метрике mAP_{0.5} с датасетом AugmentedDataset ниже, чем у датасетов, аугментированных поворотами и отражениями изображений по отдельности, хотя и немного выше, чем у YOLOv7-Tiny, обученной на оригинальном датасете (OriginDataset).

Из этого можно сделать вывод, что алгоритмы аугментации изображений с помощью поворотов и отражений подходят для повышения эффективности обучения СНС класса YOLO, в то время как аугментация с помощью мозаичного размещения не только не повышает эффективность обучения, но и снижает его, в том числе в сочетании с другими алгоритмами.

Дальнейшее исследование планируется направить на сравнительный анализ других существующих алгоритмов аугментации, а также разработку и реализацию новых, более эффективных для процесса обучения СНС класса YOLO в задаче детектирования объектов в воздушном пространстве.

Список использованных источников

1. Li Y., Zhou L., Yan H., Shan Y., Zheng C., Liu Y., Zuo X., Qiao B. Aircraft Detection for Remote Sensing Images Based on Deep Convolutional Neural Networks // Journal of Electrical and Computer Engineering. – 2021. <https://doi.org/10.1155/2021/4685644>
2. С.А. Ткачёв. С.Г. Небаба. Исследование точности детектирования объектов на изображениях с помощью нейронных сетей YOLOv3, вычисляемых на высокопроизводительных вычислительных устройствах // Сборник трудов XIX Международной научно-практической конференции студентов, аспирантов и молодых ученых МСИТ-2022. – Томск: изд-во ТПУ, 2022. – С. 49-50.
3. Wang C.-Y., Bochkovskiy A., Liao H. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv.2207.02696. – 2022.
4. С.Г. Небаба. Сверточные нейронные сети семейства YOLO для решения задачи мультиклассификации изображений // Сборник трудов XX Международной научно-практической конференции студентов, аспирантов и молодых ученых МСИТ-2023. – Томск: изд-во ТПУ, 2023. – С. 54-56.
5. Shorten C., Khoshgoufar T.M. A survey on Image Data Augmentation for Deep Learning // J Big Data, – Vol. 6. – No. 60, – 2019. <https://doi.org/10.1186/s40537-019-0197-0>
6. Hendrycks D., Mu N., Cubuk E.D., Zoph B., Gilmer J., Lakshminarayanan B. Augmix: A simple data processing method to improve robustness and uncertainty // arXiv preprint – arXiv:1912.02781. – 2019.
7. Roboflow [Электронный ресурс]. – URL: <https://roboflow.com/> (дата обращения: 29.03.2024).
8. Метрики качества ранжирования [Электронный ресурс]. – URL: <https://habr.com/ru/company/econtenta/blog/303458> (дата обращения: 29.03.2024).
9. Intersection over Union (IoU) for object detection [Электронный ресурс]. – URL: <https://pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/> (дата обращения: 29.03.2024).
10. PyTorch [Электронный ресурс]. – URL: <https://pytorch.org/> (дата обращения: 29.03.2024).