

СВЕРТОЧНЫЕ НЕЙРОННЫЕ СЕТИ СЕМЕЙСТВА YOLO ДЛЯ РЕШЕНИЯ ЗАДАЧИ МУЛЬТИКЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ

Небаба С.Г.

Томский политехнический университет, ИШИТР, ст. преподаватель, e-mail: stepanlfx@tpu.ru

Введение

Классификация летающих объектов (распознавание класса объекта) 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 и скорость классификации объектов с помощью СНС во многом зависят 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.

В статье приводятся результаты исследования эффективности различных архитектур СНС семейства YOLO [3], которые считаются одними из лучших по соотношению точности и скорости классификации объектов. Исследование ведется при детектировании на изображениях объектов 4 классов: птицы, квадрокоптеры, БПЛА самолетного типа и неизвестные объекты.

Описание датасета

В исследовании использовался подготовленный датасет с 4 классами объектов: bird (птица), uav quadcopter type (БПЛА мультироторные-квадрокоптеры), uav plane type (БПЛА самолетного типа), undefined (другие, неизвестные объекты). Общий объем датасета - 6230 изображений.

Изображения в датасете сформированы таким образом, чтобы присутствовали как примеры одиночных объектов на изображении, так и примеры множества объектов в кадре. Общее число объектов по классам следующее:

- bird – 5642;
- uav quadcopter type – 4845;
- uav plane type – 3514;
- undefined – 19809.

Датасет разделен на следующие выборки:

- обучающая – 4358 изображений (70%);
- проверочная (валидационная) – 1249 изображений (20%);
- тестовая – 623 изображений (10%).

Размер изображений в датасете 416*416 пикселей.

Описание исследуемых архитектур СНС

Для исследования было выбрано несколько архитектур СНС семейства YOLO.

yolov4-tiny – сжатая версия СНС yolov4, содержит всего 29 сверточных слоев и ряд других оптимизаций, позволяющих добиться очень высокой скорости обработки изображений, даже без использования графического процессора.

yolov4-csp – модификация архитектуры yolov4, содержащая 142 сверточных слоя и использующая механизм проброса признаков (Cross Stage Partial, CSP) между слоями для снижения сложности СНС.

yolov4-p6 – модификация yolov4, содержащая 289 сверточных слоев и настроенная на обработку больших изображений (1280*1280), однако требующая при этом существенных вычислительных ресурсов и большой объем памяти для обучения [4].

yolov7-tiny - сжатая версия СНС yolov7, которая является наиболее актуальной СНС класса YOLO, содержит всего 87 сверточных слоев и ряд других оптимизаций.

yolov7 – наиболее актуальная на данный момент СНС класса YOLO, содержит 132 сверточных слоя. Отличается от предыдущих версий СНС этого класса 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 [5].

Результаты исследования точности и скорости классификации объектов

В качестве метрики для оценки точности классификации объектов на изображениях использовалась метрика Average Precision (AP) [6]. Она рассчитывается для каждого класса объектов отдельно, следовательно, необходимо рассчитать количество значений AP равное числу рассматриваемых в поставленной задаче классов объектов. Значения AP усредняются для получения метрики mAP – среднего значения AP по всем классам. Метрика AP рассчитывается при заданном пороге Intersection over Union, IoU [7]. Использовался порог IoU=0.5.

Все результаты по точности и быстродействию исследуемых СНС были получены с помощью фреймворка Darknet, модифицированного для работы с СНС семейства YOLO [8].

Число эпох обучения у всех сетей было равным 200.

Тестирование СНС проводилось со следующими параметрами:

- доверительный порог, *conf_thres* = 0,25.
- порог IoU, *iou_thres* = 0,5.

Результаты тестирования различных архитектур СНС по метрикам AP и mAP на проверочной (валидационной) и тестовой выборках представлены в таблицах 1 и 2. Для архитектур, обучение которых проводилось на различных размерах датасета, показаны наилучшие результаты.

Таблица 1

Точность классификации объектов на изображениях проверочной выборки

Архитектура СНС	bird, AP _{0.5}	uav plane type, AP _{0.5}	uav quadcopter type, AP _{0.5}	undefined, AP _{0.5}	mAP _{0.5}
<i>yolov4-tiny</i>	0,584	0,776	0,826	0,666	0,713
<i>yolov4-csp</i>	0,726	0,877	0,897	0,848	0,837
<i>yolov4-p6</i>	0,779	0,885	0,906	0,827	0,849
<i>yolov7-tiny</i>	0,685	0,828	0,868	0,781	0,791
<i>yolov7</i>	0,767	0,888	0,903	0,847	0,851

Таблица 2

Точность классификации объектов на изображениях тестовой выборки

Архитектура СНС	bird, AP _{0.5}	uav plane type, AP _{0.5}	uav quadcopter type, AP _{0.5}	undefined, AP _{0.5}	mAP _{0.5}
<i>yolov4-tiny</i>	0,564	0,754	0,799	0,65	0,691
<i>yolov4-csp</i>	0,724	0,845	0,897	0,858	0,831
<i>yolov4-p6</i>	0,675	0,828	0,839	0,799	0,785
<i>yolov7-tiny</i>	0,649	0,804	0,856	0,777	0,771
<i>yolov7</i>	0,784	0,881	0,896	0,877	0,859

Оценка скорости вычисления СНС различной архитектуры производилась путем вычисления среднего времени обработки одного изображения на видеокарте NVIDIA Quadro RTX 6000 с объемом видеопамати 24 Гб. Результаты по оценке скорости вычисления каждой из исследуемых архитектур СНС представлены в таблице 3.

Таблица 3

Результаты оценки скорости вычисления СНС

Архитектура СНС	inference, мс	Общее время, мс
<i>yolov4-tiny</i>	5,3	5,4
<i>yolov4-csp</i>	21,4	22,9
<i>yolov4-p6</i>	87,7	89,3
<i>yolov7-tiny</i>	10,8	11
<i>yolov7</i>	20,8	22,1

Заключение

Точность детектирования и классификации объектов на изображениях с помощью этих СНС выше в среднем у более современных архитектур (*yolov7* более точная, чем *yolov4-csp*) и у архитектур с большим числом сверточных слоев (*yolov7*, *yolov4-csp* и *yolov4-pb* показывают более точные результаты, чем *yolov4-tiny* и *yolov7-tiny*). На это указывают результаты исследований как на проверочной, так и на тестовой выборках. Точность немного ниже на изображениях тестовой выборки, чем на изображениях проверочной выборки для всех архитектур (кроме *yolov7*). Так, для архитектуры *yolov4-pb* по сравнению с проверочной выборкой точность детектирования объектов в тестовой выборке по метрике $mAP_{0.5}$ снизилась на 6,5%. У остальных архитектур разница между результатами на проверочной и тестовой выборках не настолько большая.

Точность детектирования и классификации объектов класса «Птицы» по метрике $AP_{0.5}$ хуже, чем для объектов других классов. Это объясняется меньшим числом изображений объектов этого класса в датасете, малыми размерами птиц на этих изображениях, а также схожестью силуэтов птиц в некоторых ракурсах с объектами других классов.

Скорость детектирования и классификации объектов выше на архитектурах СНС с меньшим количеством сверточных слоев. Если архитектуры *yolov4-tiny* и *yolov7-tiny* можно полноценно использовать при обработке кадров в режиме реального времени (90-180 к/с), а *yolov4-csp* и *yolov7* условно тоже можно считать подходящими для решения задач реального времени (30 к/с), то архитектура *yolov4-pb*, несмотря на хорошие результаты по точности, проигрывает *yolov7* и в точности, и в скорости обработки изображений (11 к/с).

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.
2. С.А. Ткачев. 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. YOLO: Real-Time Object Detection [Электронный ресурс]. – URL: <https://pjreddie.com/darknet/yolo> (дата обращения 26.02.2023).
4. Wang C.-Y., Bochkovskiy A., Liao H. Scaled-YOLOv4: Scaling Cross Stage Partial Network. – 2021.
5. Wang C.-Y., Bochkovskiy A., Liao H. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. – 2022.
6. Метрики качества ранжирования [Электронный ресурс]. – URL: <https://habr.com/ru/company/econtenta/blog/303458> (дата обращения 26.02.2023).
7. IOU (Intersection over Union) [Электронный ресурс]. – URL: <https://medium.com/analytics-vidhya/iou-intersection-over-union-705a39e7acef> (дата обращения 26.02.2023).
8. YOLOv4 / Scaled-YOLOv4 / YOLO - Neural Networks for Object Detection (Windows and Linux version of Darknet) [Электронный ресурс]. – URL: <https://github.com/AlexeyAB/darknet> (дата обращения 26.02.2023).