

АНАЛИЗ ЭФФЕКТИВНОСТИ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ YOLOV4-CSP ПРИ РЕШЕНИИ ЗАДАЧИ МУЛЬТИКЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ

Ткачёв С.А.

Томский политехнический университет, ИШИТР, аспирант, e-mail: sat12@tpu.ru

Введение

Важнейшей задачей при контроле воздушного пространства является задача обнаружения и классификации (она же задача детектирования) летающих объектов: различных пилотируемых и беспилотных летательных аппаратов (БПЛА), птиц и т.п. Наиболее широко при решении этой задачи применяются системы детектирования объектов, работающие в режиме реального времени. В последние годы все чаще с целью детектирования таких объектов начинают создавать системы компьютерного зрения на основе современных сверточных нейронных сетей (СНС) [1].

В статье приводятся результаты исследования эффективности СНС YOLOv4-CSP при детектировании на изображениях объектов четырех классов: птицы, квадрокоптеры, БПЛА самолетного типа и неизвестные объекты.

Задача исследования эффективности СНС YOLOv4-CSP

В качестве объекта исследований была выбрана СНС перспективной архитектуры YOLOv4-CSP [2], используемая для решения задачи детектирования (одномоментного обнаружения, локализации и классификации) объектов воздушного пространства четырех классов. Отличительной чертой этой архитектуры сети класса YOLO является наличие пропусков признаков (англ. cross stage partial – CSP). Для проведения исследования был подготовлен датасет с объектами четырех классов (птицы, квадрокоптеры, БПЛА самолетного типа и неизвестные объекты), размер которого 6230 изображений. Датасет был разделен на следующие выборки: обучающая – 4358 изображений (70%); проверочная (валидационная) – 1249 изображений (20%); тестовая – 623 изображений (10%).

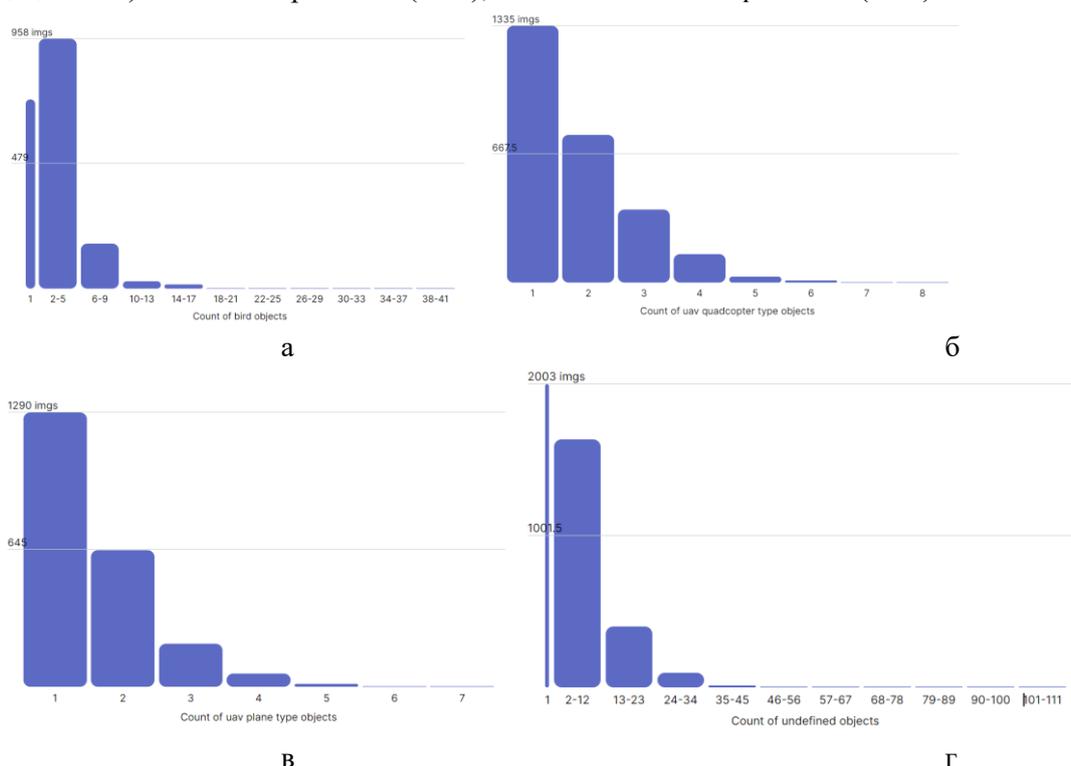


Рис. 1. Распределение числа изображений в датасете от количества объектов на отдельном изображении:

а – птицы, б – квадрокоптеры, в – БПЛА самолетного типа, г – неизвестные объекты

Распределение числа изображений в датасете в зависимости от количества объектов на отдельном изображении в случае каждого из классов объектов представлено на рисунке 1.

По оси Y отобразено количество изображений в датасете с соответствующим числом объектов, а по оси X – число объектов на отдельном изображении. При проведении исследования в качестве метрики для оценки точности детектирования (классификации) объектов на изображениях использовалась известная метрика Average Precision (AP) [3]. Она рассчитывается для каждого класса объектов. Это означает, что необходимо было получить столько значений AP, сколько в поставленной задаче рассматривается классов объектов. Эти значения AP усредняются для получения метрики mAP – среднего значения AP по всем классам. Метрика AP рассчитывается при заданном пороге IoU [4]. Использовался порог IoU = 0,5, а также метрика AP рассчитывалась для диапазона пороговых значений 0,5:0,95 с шагом 0,05. При обучении, валидации и исследовании (тестировании) СНС YOLOv4-CSP использовался фреймворк PyTorch.

Обучение СНС YOLOv4-CSP проводилось на входных изображениях трех разных размеров: 416x416, 512x512 и 608x608 пикселей в течение 200 эпох. Результаты обучения этой СНС представлены в таблице 1.

Таблица 1

Оценки точности детектирования объектов на изображениях обучающей выборки

Размер изображения, пикселей	mAP _{0.5}	mAP _{0.5:0.95}
416x416	0,883	0,645
512x512	0,898	0,674
608x608	0,910	0,695

Исследование эффективности (тестирование) СНС проводилось на входных изображениях также трех разных размеров: 416x416, 512x512 и 608x608 пикселей с параметрами batch size = 16 и confidence threshold = 0,4. Результаты исследования СНС представлены в таблицах 2 и 3.

Таблица 2

Точность детектирования объектов на изображениях проверочной и тестовой выборки для каждого класса объектов

Размер изображения, пикселей	Класс	Проверочная выборка		Тестовая выборка	
		AP _{0.5}	AP _{0.5:0.95}	AP _{0.5}	AP _{0.5:0.95}
416x416	птицы	0,690	0,546	0,677	0,532
	квадрокоптеры	0,869	0,666	0,847	0,644
	БПЛА самолетного типа	0,900	0,741	0,891	0,713
	неизвестные объекты	0,814	0,505	0,812	0,500
512x512	птицы	0,704	0,572	0,701	0,560
	квадрокоптеры	0,898	0,699	0,857	0,681
	БПЛА самолетного типа	0,913	0,765	0,892	0,732
	неизвестные объекты	0,834	0,536	0,836	0,539
608x608	птицы	0,751	0,598	0,751	0,597
	квадрокоптеры	0,879	0,712	0,856	0,698
	БПЛА самолетного типа	0,919	0,782	0,893	0,752
	неизвестные объекты	0,850	0,561	0,858	0,568

Таблица 3

Точность детектирования объектов на изображениях проверочной и тестовой выборки

Размер изображения, пикселей	Проверочная выборка		Тестовая выборка	
	mAP _{0,5}	mAP _{0,5:0,95}	mAP _{0,5}	mAP _{0,5:0,95}
416x416	0,818	0,614	0,807	0,597
512x512	0,837	0,643	0,821	0,628
608x608	0,850	0,663	0,840	0,654

Оценка скорости детектирования объектов на изображениях (оценка скорости вычисления СНС путем вычисления времени обработки одного изображения) производилась на видеокарте NVIDIA Quadro RTX 6000 с объемом видеопамати 24 ГБ. Результаты по оценке скорости вычисления СНС представлены в таблице 4.

Таблица 4

Результаты оценки скорости вычисления СНС на проверочной и тестовой выборках

Размер изображения, пикселей	Время обработки одного изображения, мс	
	Проверочная выборка	Тестовая выборка
416x416	3,1	3,1
512x512	4,4	4,5
608x608	5,9	5,9

Заключение

Сравнительный анализ всех полученных результатов по оценке точности и скорости детектирования объектов на изображениях с помощью СНС YOLOv4-CSP позволил сделать следующие выводы.

Точность детектирования объектов на изображениях с помощью этой СНС выше на входных изображениях большего размера. Это имеет место для каждой из выборок датасета. Например, при изменении размера изображения с 416x416 на 608x608 пикселей на тестовой выборке точность детектирования объектов по метрике mAP_{0,5} выросла на 3,3%, а по метрике mAP_{0,5:0,95} – на 5,7%.

Точность детектирования объектов ниже на изображениях тестовой выборки, чем на изображениях обучающей и проверочной выборок. Так, на изображениях с размером 608x608 пикселей по сравнению с обучающей выборкой точность детектирования объектов по метрике mAP_{0,5} снизилась на 7%, а по метрике mAP_{0,5:0,95} – на 4,1%.

Точность детектирования объектов класса «птицы» по метрике AP_{0,5} хуже, чем для объектов других классов. Это объясняется меньшим числом изображений объектов этого класса в датасете, а также, в основном, малыми размерами птиц на этих изображениях. По метрике AP_{0,5:0,95} точность ниже при детектировании объектов класса «неизвестные объекты».

Скорость детектирования объектов выше на изображениях меньшего размера. Так, при изменении размера изображения с 416x416 на 608x608 пикселей время обработки одного изображения увеличилось на 2,8 мс.

Список использованных источников

1. Ткачёв С.А., Небаба С.Г. 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. – С. 49–50.
2. Scaled YOLO v4 [Электронный ресурс]. – URL: <https://alexeyab84.medium.com/scaled-yolo-v4-is-the-best-neural-network-for-object-detection-on-ms-coco-dataset-39dfa22fa982> (дата обращения 23.02.2023).
3. Метрики качества ранжирования [Электронный ресурс]. – URL: <https://habr.com/ru/company/econtenta/blog/303458> (дата обращения 23.02.2023).
4. IOU (Intersection over Union) [Электронный ресурс]. – URL: <https://medium.com/analytics-vidhya/iou-intersection-over-union-705a39e7acef> (дата обращения 23.02.2023).