

ДЕТЕКТИРОВАНИЕ ОБЪЕКТОВ ВОЗДУШНОГО ПРОСТРАНСТВА НА ИЗОБРАЖЕНИЯХ С ИСПОЛЬЗОВАНИЕМ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ YOLOV5

Клековкин В.А.

*Томский политехнический университет, ИШИТР,
группа А2-39, e-mail: vak37@tpu.ru*

Аннотация

Рассматривается задача детектирования (обнаружения и классификации) летающих объектов трех классов на изображениях с помощью моделей сверточных нейронных сетей (СНС) YOLOv5s и YOLOv5x. Для обучения, валидации и исследования эффективности этих моделей разработан датасет. Он содержит размеченные изображения с летающими объектами трех классов: Беспилотный летательный аппарат (БПЛА) самолетного типа, включая «летающее крыло», БПЛА вертолетного типа и Птица. Исследования предложенных моделей СНС по точности классификации таких объектов на изображениях тестовой выборки датасета показали, что обе модели показывают высокие и весьма близкие результаты по всем используемым метрикам оценки точности классификации. В результате исследований также получено, что по скорости вычислений модель YOLOv5s значительно превосходит модель YOLOv5x.

Ключевые слова: сверточная нейронная сеть YOLO, датасет, беспилотный летательный аппарат, птица, точность классификации летающих объектов

Введение

В настоящее время практически важными являются задачи мониторинга воздушного пространства с целью выявления и классификации в нем летающих объектов. В числе таких объектов в воздушном пространстве рассматриваются птицы, беспилотные летательные аппараты (БПЛА) различных типов (самолётного типа, в том числе «летающее крыло» и вертолетного типа), пилотируемые летательные аппараты (самолеты, вертолеты, дельтапланы и планеры). Несмотря на то, что задачи обнаружения и распознавания подвижных объектов воздушного пространства на изображениях изучаются достаточно давно, исследования на эту тему и на сегодняшний день продолжают оставаться актуальными [1]. При этом для детектирования (обнаружения и классификации) таких объектов перспективными являются модели класса YOLO [2].

Данная работа посвящена исследованию двух моделей из класса YOLO на созданном для этих целей датасете в части точности классификации и скорости обнаружения летающих объектов трех классов.

Постановка задачи

Анализ сетей класса YOLO показывает, что наиболее перспективными являются модели семейства YOLOv5 [2]. Будем рассматривать подвижные объекты в воздушном пространстве трех классов: БПЛА самолетного типа, включая «летающее крыло», БПЛА вертолетного типа и Птица. Целью работы является проведение исследований моделей YOLOv5s и YOLOv5x из этого семейства на датасете, содержащем изображения летающих объектов таких трех классов. Изображения в датасете могут содержать один или более объектов, причем это могут быть объекты как одного, так и нескольких классов. Объекты на изображениях датасета будут иметь разные размеры (разномасштабные объекты), при этом размеры их находятся в пределах от 20x20 до 128x128 пикселей.

Для реализации поставленной цели необходимо сформировать датасет, обучить и валидировать модели YOLOv5s и YOLOv5x на его обучающих и валидационных выборках, а также исследовать эти модели на тестовом наборе данных датасета по точности классификации летающих объектов и скорости вычисления моделей.

Формирование датасета

При создании датасета для обучения, валидации и тестирования предложенных моделей СНС осуществлялся поиск изображений с летающими объектами трех классов: БПЛА самолетного типа, включая «летающее крыло», БПЛА вертолетного типа и Птица. Подбор изображений выполнялся в соответствии со следующим требованием: исходные изображения должны максимально реалистично представлять встречающиеся сценарии расположения объектов в воздушном пространстве (наличие на

изображении одного, двух и более летающих объектов перечисленных классов, наличие фона в виде участков земной поверхности или облаков и т. д.). Также часть полученных изображений была подвергнута аугментации (создание дополнительных изображений из имеющихся данных) для увеличения объема выборок. Пример исходных изображений для формируемого датасета представлен на рис. 1.

Далее все изображения были приведены к единому размеру (416x416 пикселей) и размечены с помощью средств системы Roboflow [3]. Часть размеченных изображений из датасета в качестве примера представлена на рис. 2.



Рис. 1. Пример исходных изображений для формируемого датасета

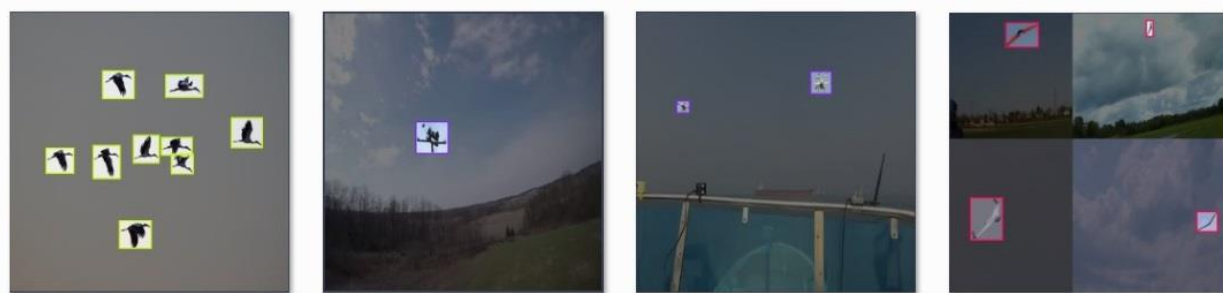


Рис. 2. Пример размеченных изображений из сформированного датасета

В итоге все размеченные изображения датасета, содержащего 5235 объектов, были разделены физически на три выборки: обучающая (70 % объема датасета), валидационная (20 %) и тестовая (10 %). На рис. 3 представлены количественные характеристики датасета по классам летающих объектов.

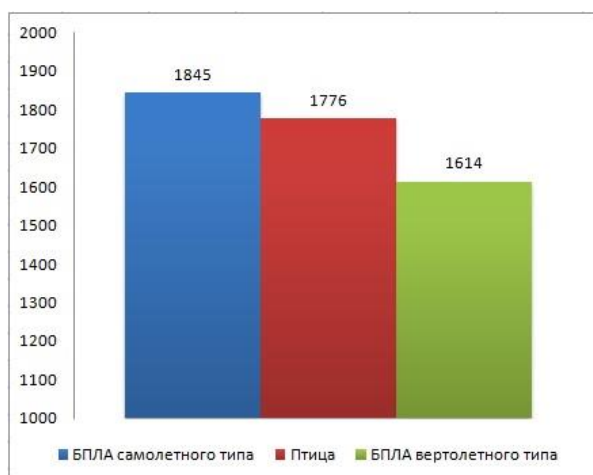


Рис. 3. Количество объектов по классам в датасете

Обучение и исследование эффективности моделей YOLOv5s и YOLOv5x

Программная реализация предложенной модели СНС осуществлялась на языке программирования Python версии 3.6 с использованием библиотеки PyTorch [4]. Для оценки качества моделей СНС (точности классификации объектов) использовались известные метрики Precision, Recall,

Average Precision (AP) 0.5, mean Average Precision (mAP) 0.5, Average Precision (AP) 0.5-0.95 и mean Average Precision (mAP) 0.5-0.95 [5, 6]. Эксперименты проводились на компьютере с графическим процессором Tesla T4, 15102MiB.

Был проведен ряд экспериментов при обучении моделей с изменением гиперпараметра `batch_size`, а также с изменением количества эпох [7]. В таблице 1 показаны результаты по точности классификации при валидации моделей в зависимости от размера `batch_size`. Было выявлено, что лучшие результаты по метрикам достигаются для каждой из моделей при `batch_size` = 8 и количестве эпох, равном 120. На рисунках 4 и 5 приведены графики обучения моделей YOLOv5s и YOLOv5x, отображающие результаты по метрике mAP 0.5 в зависимости от количества эпох, `batch_size`=8. Можно увидеть, что эффект плато достигается при 100-120 эпохах обучения для каждой модели, поэтому при обучении выбрано количество эпох, равное 120.

Таблица 1

Точность классификации при валидации моделей в зависимости от `batch_size`

batch_size	mAP, YOLOv5s	mAP, YOLOv5x
8	0,953	0,954
16	0,935	0,947
32	0,929	0,931
48	0,912	0,915

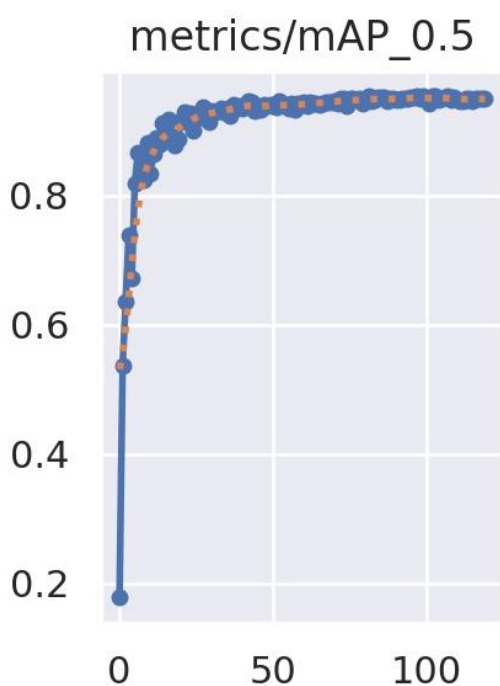


Рис. 4. Зависимость значения метрики mAP 0.5 от количества эпох обучения для модели YOLOv5s, `batch_size`=8



Рис. 5. Зависимость значения метрики mAP 0.5 от количества эпох обучения для модели YOLOv5x, `batch_size`=8

Результаты обучения, валидации и тестирования моделей на соответствующих выборках из датасета представлены в таблицах 2, 3, 4.

Таблица 2

Результаты обучения моделей

Класс	Precision		Recall		AP0.5, mAP0.5		AP0.5-0.95, mAP0.5-0.95	
	YOLOv5s	YOLOv5x	YOLOv5s	YOLOv5x	YOLOv5s	YOLOv5x	YOLOv5s	YOLOv5x
БПЛА самолетного типа	0,97	0,992	0,992	0,997	0,988	0,995	0,833	0,889
Птица	0,963	0,996	0,983	0,99	0,985	0,995	0,86	0,921
БПЛА вертолетного типа	0,963	0,979	0,987	0,999	0,989	0,993	0,833	0,884
все классы	0,965	0,989	0,987	0,995	0,987	0,994	0,842	0,898

Таблица 3

Результаты валидации моделей

Класс	Precision		Recall		AP0.5, mAP0.5		AP0.5-0.95, mAP0.5-0.95	
	YOLOv5s	YOLOv5x	YOLOv5s	YOLOv5x	YOLOv5s	YOLOv5x	YOLOv5s	YOLOv5x
БПЛА самолетного типа	0,928	0,958	0,96	0,951	0,975	0,973	0,756	0,771
Птица	0,847	0,915	0,94	0,956	0,919	0,938	0,76	0,792
БПЛА вертолетного типа	0,917	0,903	0,952	0,952	0,966	0,952	0,777	0,779
Все классы	0,897	0,925	0,951	0,953	0,953	0,954	0,764	0,781

Таблица 4

Результаты исследования моделей на тестовых данных

Класс	Precision		Recall		AP0.5, mAP0.5		AP0.5-0.95, mAP0.5-0.95	
	YOLOv5s	YOLOv5x	YOLOv5s	YOLOv5x	YOLOv5s	YOLOv5x	YOLOv5s	YOLOv5x
БПЛА самолетного типа	0,921	0,943	0,941	0,954	0,948	0,943	0,718	0,724
Птица	0,852	0,892	0,903	0,928	0,9	0,887	0,714	0,723
БПЛА вертолетного типа	0,894	0,903	0,98	0,964	0,965	0,954	0,761	0,777
Все классы	0,889	0,913	0,941	0,948	0,938	0,928	0,731	0,741

На рис. 6 приведен пример детектирования объекта на изображении из тестового набора данных.

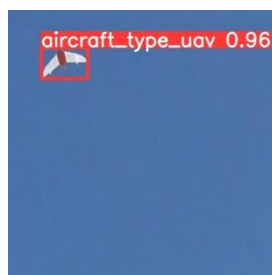


Рис. 6. Пример детектирования объекта на изображении из тестового набора данных

Исходя из результатов исследований, можно сделать вывод, что обе модели успешно справляются с классификацией объектов на изображениях и показывают весьма близкие результаты по всем используемым метрикам оценки точности классификации.

Таблица 5 содержит результаты исследования моделей по скорости вычислений. Для оценки скорости вычислений в исследованиях выбраны следующие показатели [8, 9]:

- inference – время, необходимое для вывода одного изображения;
- NMS – время постобработки одного изображения с помощью метода для устранения повторяющихся обнаружений и выбора наиболее релевантных обнаруженных объектов; это способствует уменьшению количества ложных срабатываний и вычислительной сложности алгоритма обнаружения;
- FPS – количество обработанных изображений в секунду;
- суммарное время обработки всех изображений из тестового набора данных (количество изображений тестового набора – 396).

Таблица 5

Результаты исследования моделей по скорости вычислений

Модель	Inference, мс	NMS, мс	FPS	Суммарное время обработки всех изображений из тестового набора данных, с
YOLOv5s	5,3	5,6	91	4,3
YOLOv5x	11	6,7	56	7

Исходя из результатов исследований, можно сделать вывод, что модель YOLOv5s значительно лучше модели YOLOv5x по каждому из используемых показателей скорости вычислений.

Заключение

Для решения задачи исследования моделей YOLOv5s и YOLOv5x по скорости вычислений и точности классификации был разработан датасет, содержащий изображения с летающими объектами трех классов: БПЛА самолетного типа, включая «летающее крыло», БПЛА вертолетного типа и Птица. Эти модели обучались и проверялись, а также исследовались, соответственно, на обучающей, валидационной и тестовой выборках созданного датасета. Исследования данных моделей СНС по точности классификации летающих объектов на изображениях тестовой выборки датасета показали, что обе модели успешно справляются с поставленной задачей детектирования таких объектов и по всем метрикам точности классификации показывают близкие результаты. Однако модель YOLOv5s значительно лучше модели YOLOv5x по скорости вычислений.

Список использованных источников

1. Nebaba S. G., Markov N. G. Effectiveness of Moving Objects Detecting and Tracking in Airspace by Images in Nearinfrared // Light & Engineering. – 2022. – Vol. – 30. – No 2. – P. 62–69.
2. Nelson J. Your Comprehensive Guide to the YOLO Family of Models [Электронный ресурс]. – URL: <https://blog.roboflow.com/guide-to-yolo-models/> (дата обращения: 10.03.2024).
3. Roboflow [Электронный ресурс]. – URL: <https://roboflow.com/> (дата обращения: 12.03.2024).
4. Pytorch [Электронный ресурс]. – URL: <https://pytorch.org/> (дата обращения: 12.03.2024).

5. Метрики в задачах машинного обучения [Электронный ресурс]. – URL: <https://habr.com/ru/company/ods/blog/328372/> (дата обращения: 12.03.2024)
6. mAP (mean Average Precision) в детекции объектов [Электронный ресурс]. – URL: <https://ml.i-neti.ru/map-mean-average-precision/> (дата обращения: 14.03.2024).
7. Difference Between a Batch and an Epoch in a Neural Network [Электронный ресурс]. – URL: <https://machinelearningmastery.com/difference-between-a-batch-and-an-epoch/> (дата обращения: 16.03.2024).
8. Object Detection: Non-Maximum Suppression (NMS) [Электронный ресурс]. – URL: <https://kikaben.com/object-detection-non-maximum-suppression/> (дата обращения: 18.03.2024).
9. Сравнение производительности моделей обнаружения объектов YOLO v5, v6 и v7 [Электронный ресурс]. – URL: <https://habr.com/ru/articles/702790/> (дата обращения: 18.03.2024).