

ДЕТЕКТИРОВАНИЕ ЛЕТАЮЩИХ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ С ПОМОЩЬЮ СВЁРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ СЕМЕЙСТВА YOLO

Мунько А.С.¹, Марков Н.Г.²

¹ Томский политехнический университет, ИШИТР, магистрант гр. 8ВМ1И, e-mail: asm104@tpu.ru

² Томский политехнический университет, ИШИТР, профессор, e-mail: markovng@tpu.ru

Введение

В настоящее время во всем мире интенсивно разрабатывается актуальная научно-техническая проблема создания интеллектуальных систем компьютерного зрения (СКЗ) различного назначения. Такие системы сегодня востребованы для решения многих прикладных задач, в том числе в различных отраслях промышленности. Особое научное и практическое значение имеют мобильные (возимые или носимые) СКЗ на основе современных моделей свёрточных нейронных сетей (СНС). Вычислительные возможности мобильных СКЗ обычно весьма ограничены, что накладывает жёсткие требования на алгоритмы предварительной обработки изображений на входе СКЗ и на модели СНС, используемые для анализа таких изображений с целью распознавания на них объектов различной физической природы.

В данной работе исследуется семейство из трех моделей СНС класса YOLO для детектирования (обнаружения, локализации и классификации) летающих объектов на изображениях. Результаты исследований позволяют выбрать наиболее эффективную модель СНС из этого семейства для создания мобильных СКЗ реального времени.

Постановка задачи исследования семейства моделей СНС класса YOLO

В работе ставится задача исследования моделей СНС класса YOLO, позволяющих детектировать объекты на RGB-изображениях четырёх классов: птица (птицы), беспилотные летательные аппараты (БПЛА) самолётного типа, БПЛА вертолётного типа и неизвестные объекты. При создании мобильных СКЗ важно, чтобы реализованные в них модели СНС удовлетворяли двум основным требованиям: классификация объектов на изображениях должна проводиться с высокой точностью, а процесс детектирования объектов в таких СКЗ — выполняться в масштабе реального времени (анализ изображений с помощью моделей СНС должен выполняться в темпе их поступления на вход СКЗ). Учитывая весьма высокую сложность моделей СНС класса YOLO, обеспечивающих точность классификации изображений по известным метрикам не хуже 0,9, и обычно скромные вычислительные ресурсы мобильных СКЗ, необходимо искать модели СНС, одновременно удовлетворяющие этим требованиям.

Проведённый анализ моделей класса YOLO показал, что наиболее перспективными для достижения компромисса при удовлетворении таким требованиям является семейство моделей YOLOv4, YOLOv5 и YOLOv7 [3, 4]. Более того, исследованию и выявлению наиболее эффективной из них подлежат компактные (малое число свёрточных слоёв) модели СНС из этого семейства: YOLOv4 Tiny, YOLOv5 Small и YOLOv7 Tiny. Обучение, валидация и исследование первой из этих моделей проводились с использованием фреймворка Darknet, а второй и третьей моделей — с использованием фреймворка PyTorch.

Формирование датасета

Исходными данными для формирования датасета являются изображения размером 416×416 пикселей объектов четырёх классов: птица (птицы), БПЛА самолётного типа, БПЛА вертолётного типа и неизвестные объекты. Создание датасета осуществлялось путём объединения существующих датасетов и дополнительного сбора и разметки изображений с помощью системы Roboflow, содержащих объекты перечисленных классов.

Полученный датасет включает в себя 15 235 исходных изображений, содержащих 38 553 объектов, и столько же файлов аннотаций. Изображения и аннотации датасета содержат один и более объектов, в том числе разных классов. Распределение в виде столбчатых диаграмм числа изображений в датасете в зависимости от количества объектов на каждом изображении, входящем в тот или иной класс, показано на рисунке 1. По оси Y приведено количество изображений каждого класса в датасете с соответствующим числом объектов, а по оси X — число объектов на отдельном изображении.

Датасет был разделён на обучающую, валидационную и тестовую выборки, объёмы которых составляют, соответственно, 80%, 10% и 10% от всего объёма изображений в датасете.

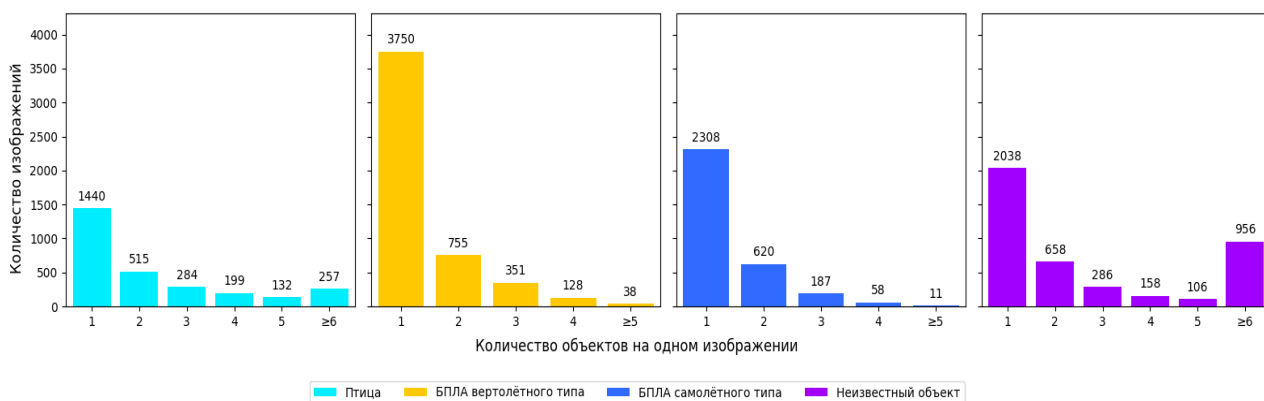


Рис. 1. Распределение числа изображений в датасете по классам в зависимости от количества объектов на отдельном изображении

Обучение и исследование моделей СНС из выбранного семейства

Обучение, валидация и исследование первой модели YOLOv4 Tiny из выбранного семейства проводилось с использованием фреймворка Darknet, а второй и третьей моделей YOLOv5 Small и YOLOv7 Tiny – с использованием фреймворка PyTorch. Применение разных фреймворков не позволяет задать полностью идентичные параметры при обучении этих моделей СНС. Однако схожими и важными являются параметры:

- количество итераций для модели YOLOv4 Tiny – 16 000;
- количество эпох для моделей YOLOv5 Small и YOLOv7 Tiny – 100.

Эпоха – это процесс прохождения всего набора данных через нейронную сеть вперед и назад. Итерации можно перевести в эпохи: при использовании размера пакета *Batch*, равного 64, на одну эпоху потребуется 191 итерация. При обучении модели YOLOv4 Tiny выполняется 16 000 итераций или 85 эпох, что близко к количеству эпох при обучении двух других моделей.

Для оценки точности детектирования (точности классификации) объектов каждого класса на изображениях использовались известные метрики Precision, Recall, F1, AP_{0,5} и AP_{0,5:0,95}. Метрика AP рассчитывалась при заданном пороге IoU [2]. Использовался порог 0,5, а также диапазон пороговых значений 0,5:0,95 с шагом 0,05. Кроме того, использовалась метрика mAP – среднее значение AP по всем четырём классам объектов. Во время обучения моделей СНС проводилась их валидация на валидационной выборке, результаты которой также оценивались с помощью этих метрик.

Исследование каждой из обученных трех моделей СНС проводилось с использованием тестовой выборки из 1 526 изображений. Примеры классификации объектов с помощью модели YOLOv5 Small приведены на рисунке 2. Видно, что модель точно локализует прямоугольниками и правильно классифицирует один или несколько объектов на изображении.

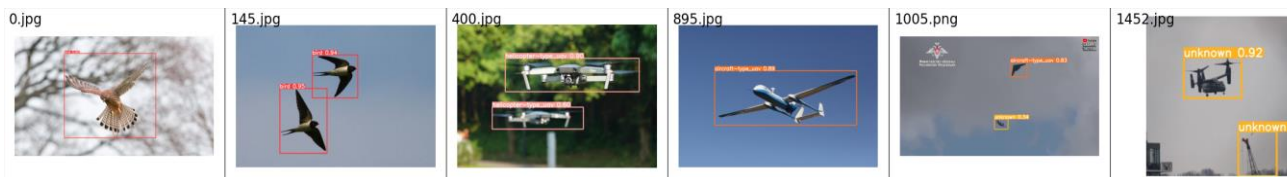


Рис. 2. Классификация объектов на изображениях с помощью модели YOLOv5 Small

Результаты исследования моделей СНС по точности детектирования (классификации) объектов на изображениях для различных метрик представлены в таблице 1. Из неё следует, что по точности классификации объектов предпочтение следует отдать модели YOLOv5 Small.

Таблица 1

Точность классификации моделей СНС на тестовой выборке изображений

Модель СНС	Класс	Precision	Recall	F1	AP _{0,5} ; mAP _{0,5}	AP _{0,5:0,95} ; mAP _{0,5:0,95}
	Птица	—	—	—	0,457	—

Модель СНС	Класс	Precision	Recall	F1	AP _{0,5} ; mAP _{0,5}	AP _{0,5:0,95} ; mAP _{0,5:0,95}
YOLOv4 Tiny	БПЛА вертолётного типа	—	—	—	0,756	—
	БПЛА самолётного типа	—	—	—	0,690	—
	Неизвестный объект	—	—	—	0,489	—
	Все классы	0,530	0,490	0,510	0,600	—
YOLOv5 Small	Птица	0,936	0,763	0,840	0,840	0,687
	БПЛА вертолётного типа	0,945	0,969	0,957	0,970	0,789
	БПЛА самолётного типа	0,937	0,933	0,935	0,966	0,784
	Неизвестный объект	0,912	0,836	0,872	0,902	0,597
	Все классы	0,932	0,875	0,902	0,920	0,714
YOLOv7 Tiny	Птица	0,912	0,669	0,772	0,751	0,555
	БПЛА вертолётного типа	0,928	0,901	0,914	0,945	0,688
	БПЛА самолётного типа	0,932	0,847	0,888	0,922	0,641
	Неизвестный объект	0,848	0,710	0,773	0,802	0,463
	Все классы	0,905	0,782	0,839	0,855	0,587

Проведены исследования скорости вычисления этих моделей СНС путём оценки времени детектирования (классификации) объектов на одном изображении. Результаты в таблице 2 являются средними значениями времени классификации с учётом объёма тестовой выборки.

Таблица 2

Среднее время классификации объектов на одном изображении

Модель СНС	Inference, мс	NMS, мс	Всего, мс
YOLOv4 Tiny	—	—	11,3
YOLOv5 Small	6,9	3,3	10,5
YOLOv7 Tiny	4,9	2,6	7,5

Отметим, что при обучении, валидации и исследованиях моделей СНС использовался компьютер с процессором Intel Core i7-6700K; с ОЗУ DDR4, 32 Гб, 1333 МГц и с графическим процессором NVIDIA GeForce GTX 960, 4 Гб.

Заключение

Были выбраны, обучены и исследованы три компактные модели СНС из семейства моделей YOLO, наиболее перспективного для достижения компромисса при удовлетворении требованиям со стороны мобильной СКЗ. Результаты исследований показали, что модель YOLOv5 Small можно рекомендовать к реализации в мобильных СКЗ реального времени.

Список используемых источников

1. Rawat K. Yolo Algorithm (The Layman's Approach) [Электронный ресурс]. – URL: <https://medium.com/@kartike47/yolo-algorithm-the-laymans-approach-57766a0bec71>. (дата обращения: 20.02.2023).
2. Инфосистемы Джет. Object Detection. Распознавай и властвуй. Часть 1 [Электронный ресурс]. – URL: <https://habr.com/ru/company/jetinfosystems/blog/498294/> (дата обращения: 20.02.2023).
3. CVAT. Using YOLO for annotation in CVAT [Электронный ресурс]. – URL: <https://www.cvat.ai/post/yolo>. (дата обращения: 20.02.2023).
4. Nelson J. Your Comprehensive Guide to the YOLO Family of Models [Электронный ресурс]. – URL: <https://blog.roboflow.com/guide-to-yolo-models/> (дата обращения: 20.02.2023).