

АНАЛИЗ ЭФФЕКТИВНОСТИ КОМПАКТНЫХ МОДЕЛЕЙ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ КЛАССА YOLO ПРИ РЕШЕНИИ ЗАДАЧИ КЛАССИФИКАЦИИ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ

Беляев С.И.

Томский политехнический университет, ИШИТР, студент гр. 8K92, e-mail: sib5@tpu.ru

Введение

Компьютерное зрение — это одна из областей искусственного интеллекта, которая обучает компьютеры и позволяет им **понимать** визуальный мир. Электронно-вычислительные машины могут использовать цифровые изображения и модели глубокого обучения для точной идентификации и классификации объектов и реагирования на них.

В последние годы технологии компьютерного зрения нашли широкое распространение в различных областях науки и техники, начиная с сельскохозяйственной отрасли и заканчивая военно-стратегическими целями их применения [1]. Исходя из этого широкого круга применений, сегодня актуальны исследования эффективности моделей сверточных нейронных сетей (СНС), являющихся основой многих систем компьютерного зрения (СКЗ). Чаще всего в качестве критериев эффективности модели СНС выступают скорость ее выполнения и точность решения с ее помощью задачи классификации. Кроме того, немаловажным критерием в определении эффективности модели является оценка ее компактности - объема весовых коэффициентов сверточных слоев модели.

Целью данной работы является анализ современных моделей СНС класса YOLO, выбор из них трех наиболее компактных моделей СНС и проведение исследования эффективности этих моделей с учетом указанных критериев при решении задачи классификации объектов четырех классов на изображениях.

Постановка задачи классификации и выбор моделей СНС для ее решения

Ставится задача классификации на изображениях объектов четырех классов: беспилотные летательные аппараты (БПЛА) вертолетного типа, БПЛА самолетного типа, птицы и класс неизвестных объектов. Классификацию таких объектов необходимо выполнять с помощью современных моделей СНС. Критериями выбора таких моделей являются: высокие точность классификации и скорость выполнения СНС на вычислительных устройствах, а также требование малого объема весовых коэффициентов слоев модели, хранимых в памяти этого устройства.

Одним из классов СНС, довольно успешно справляющихся с задачей классификации объектов на изображениях, является класс моделей YOLO [2]. СНС этого класса позволяют практически в один проход (часто говорят, одномоментно) обнаружить и локализовать объекты на изображениях, а также добиться весьма высокой точности классификации таких объектов при сравнительно большой скорости выполнения всех перечисленных операций [3]. По мнению ряда специалистов, входящие в класс YOLO модели YOLOv4 Scaled, YOLOv5n и YOLO Nano [3, 4] являются не только компактными (имеют малое число сверточных слоев), но и считаются одними из лучших с точки зрения высокой скорости вычислений и точности классификации. Эти модели СНС и подлежат исследованию на эффективность путем решения с их помощью задачи классификации на изображениях объектов четырех классов.

Формирование датасета и обучение моделей СНС

Для обучения и дальнейшего исследования (тестирования обученных моделей) выбранных трех моделей СНС был сформирован из открытых источников датасет с изображениями объектов четырех классов: БПЛА вертолетного типа, БПЛА самолетного типа, птицы и класс неизвестных объектов. Изображения из датасета имеют размер 416x416 пикселей. Всего в датасете содержится 4180 изображений, из них 2900 в обучающей выборке, 800 в валидационной выборке и 480 в тестовой. К ряду изображений датасета применена мозаичная аугментация, которая позволяет улучшить способность СНС распознавать малоразмерные объекты путем помещения нескольких объектов в одно изображение. Для оценки точности классификации объектов с помощью моделей на обучающей, валидационной и тестовой выборке использовались метрики Precision, Recall и mAP_{0.5} – среднее значение известной метрики AP_{0.5} по всем четырём классам объектов.

Обучение и исследование моделей СНС проводилось в среде для проведения облачных вычислений Google Collab на GPU Tesla T4 15110. На рисунке 1 в качестве примера приведены результаты обучения модели YOLOv4 Scaled.

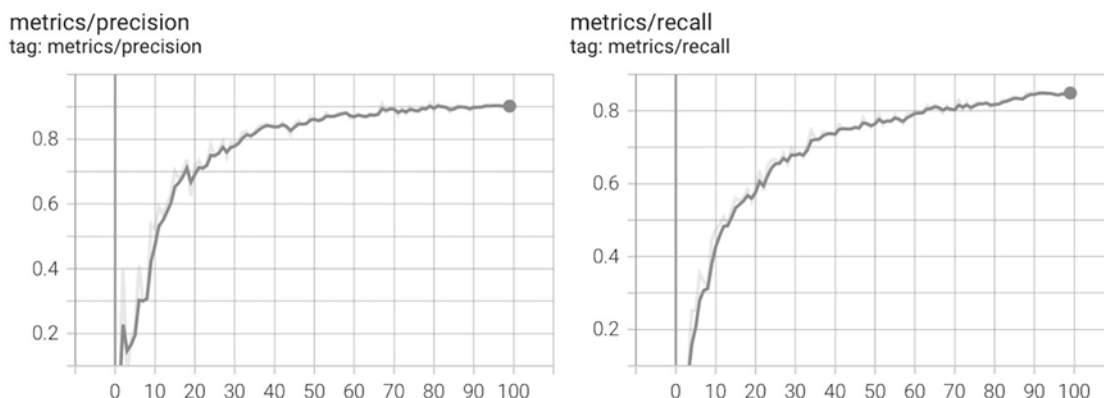


Рис. 1. Значения метрик Precision и Recall для модели YOLOv4 Scaled в зависимости от числа эпох

Результаты исследований

На тестовой выборке проведены исследования точности классификации и производительности каждой из моделей YOLOv4 Scaled и YOLOv5n и YOLO Nano с целью последующего анализа результатов и определения наиболее эффективной из них. На рисунке 2 показан пример классификации изображения с объектами двух классов: птица и БПЛА вертолетного типа. Наряду с локализацией объектов прямоугольниками показаны также коэффициенты уверенности обнаружения этих объектов. Основные результаты исследования приведены в таблице 1.



Рис. 2. Пример классификации объектов «БПЛА вертолетного типа» и «птицы» на изображении

Таблица 1

Основные результаты исследования моделей СНС на тестовой выборке

	YOLOv4 Scaled	YOLOv5n	YOLO Nano
Скорость предобработки одного изображения, мс	0,56	0,3	0,8
Скорость классификации одного изображения, мс	1,3	0,8	2,1
Скорость классификации изображений всей тестовой выборки, мс	486	288	840
Суммарное число весовых коэффициентов модели СНС, млн	10	7,5	1.8
mAP _{0.5} val	0,9	0,92	0,55
mAP _{0.5} test	0,87	0,90	0,51

Метрика $mAP_{0.5} val$ получена на валидационной выборке, а метрика $mAP_{0.5} test$ - на тестовой выборке. Нетрудно видеть, что по критериям точности классификации и скорости вычисления моделей предпочтение следует отдать модели YOLOv5n. Однако эта модель имеет больший объем весовых коэффициентов, чем модель YOLO Nano.

Заключение

Был проведен анализ эффективности трех компактных моделей СНС YOLOv4 и YOLOv5 и YOLO Nano, входящих в класс СНС YOLO. Для этого модели были обучены и исследованы на сформированном датасете, содержащем изображения с объектами четырех классов. Показано, что по критериям точности классификации и скорости вычисления модели предпочтение следует отдать модели YOLOv5n. Однако она имеет больший объем весовых коэффициентов, чем модель YOLO Nano. Для практической реализации в СКЗ можно рекомендовать модель YOLOv5n.

Список использованных источников

1. Brotherton T., Johnson T. Anomaly detection for advanced military aircraft using neural networks //2001 IEEE Aerospace Conference Proceedings (Cat. No. 01TH8542). – 2001. – Т. 6. – P. 3113–3123.
2. Redmon J. et al. You only look once: Unified, real-time object detection //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – P. 779–788.
3. Lee Y. H., Kim Y. Comparison of CNN and YOLO for Object Detection //Journal of the semiconductor & display technology. – 2020. – Т. 19. – №. 1. – С. 85–92.
4. Wong A. et al. YOLO nano: A highly compact you only look once convolutional neural network for object detection //2019 Fifth Workshop on Energy Efficient Machine Learning and Cognitive Computing-NeurIPS Edition (EMCC2-NIPS). – IEEE, 2019. – P. 22–25.