МОДЕЛИ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ YOLO ДЛЯ ДЕТЕКТИРОВАНИЯ НА ИЗОБРАЖЕНИЯХ ОБЪЕКТОВ ВОЗДУШНОГО ПРОСТРАНСТВА

Тятюшкин И.М. [1], *Кравченко А.А.* [2]
[1] *НИ ТПУ, ИШИТР, гр. 8И11, студент, imt12@tpu.ru,*
[2] *НИ ТПУ, ИШИТР, гр. 8И11 студент, aak310@tpu.ru*

Аннотация

В работе проведены комплексные исследования и выполнен сравнительный анализ полученных результатов для моделей сверточных нейронных сетей семейств YOLOv5 и YOLOv7, входящих в класс YOLO. Исследования проведены при решении задачи объектного детектирования на изображениях летающих объектов четырех классов.

Ключевые слова: сверточные нейронные сети YOLO, объектное детектирование на изображениях летающих объектов.

Введение

Сверточные нейронные сети (СНС) стали неотъемлемой частью современных систем компьютерного зрения, благодаря своей способности эффективно обрабатывать и анализировать изображения. Одной из наиболее известных и широко используемых моделей СНС являются модели класса YOLO (англ. You Only Look Once), которые выделяются своей высокой скоростью вычислений и довольно высокой точностью при решении задач объектного детектирования (обнаружения, позиционирования и классификации) на изображениях объектов различной физической природы [1, 2]. Однако получены только первые результаты комплексных исследований отдельных моделей из класса YOLO в случае детектирования на изображениях летающих объектов (ЛО) воздушного пространства. Это указывает на актуальность проведения таких исследований современных моделей YOLO с использованием соответствующих датасетов [3].

В данной статье рассмотрены модели СНС семейств YOLOv5 и YOLOv7 класса YOLO. Комплексные исследования этих моделей проведены при решении задачи объектного детектирования на изображениях объектов воздушного пространства четырех классов.

Выбор моделей СНС класса YOLO для проведения исследований

Анализ показывает, что среди наиболее известных и широко используемых моделей СНС для решения задач объектного детектирования на изображениях выделяют модели класса YOLO [2]. Они имеют высокую скорость вычислений и довольно высокую точностью детектирования объектов на изображениях. Однако появились только отдельные результаты исследований первых моделей из класса YOLO в случае детектирования объектов воздушного пространства на изображениях. Поэтому сегодня актуально проведение комплексных исследований более современных моделей YOLO. При этом имеет смысл рассматривать и исследовать эффективность моделей СНС класса YOLO с наиболее современными архитектурами, начиная с моделей семейства YOLOv5 [4,5] и заканчивая уже известными по ряду решенных прикладных задач моделями семейства YOLOv7 [6,7]. Модели семейств YOLOv5 и YOLOv7 представляют собой значительные улучшения первых моделей СНС класса YOLO, причем каждая из них имеет свои уникальные особенности и преимущества. Кратко рассмотрим особенности моделей этих семейств.

Архитектура и особенности моделей семейства YOLOv5. Пять моделей этого семейства обладают модульной архитектурой, что позволяет легко настраивать и оптимизировать их под конкретные задачи [4,5]. Отличительной чертой этих моделей является высокая скорость обработки и анализа изображений, что делает их подходящей для задач, требующих решения в реальном времени. Модели YOLOv5 используют различные оптимизации, такие как CSP

(англ. Cross Stage Partial) блоки, которые улучшают производительность моделей без значительного увеличения вычислительных затрат. Следует отметить простоту использования: модели YOLOv5 предоставляет удобные инструменты и библиотеки для обучения и развертывания моделей, что упрощает их интеграцию в существующие системы. Модели YOLOv5 широко используются в задачах, где требуется баланс между скоростью и точностью детектирования объектов на изображениях.

Архитектура и особенности моделей семейства YOLOv7. При создании архитектуры моделей этого семейства использованы передовые методы, такие как расширенные сверточные слои и улучшенные механизмы обучения. Это позволяет моделям достигать высокой точности детектирования даже на сложных наборах данных. Модели показывают повышенную точность, поскольку основное внимание разработчики уделили улучшению точности детектирования объектов. Для этого используются сложные архитектурные решения [6, 7], такие как E-ELAN (англ. Extended Efficient Layer Aggregation Network), которые позволяют более эффективно извлекать и агрегировать признаки объектов на изображениях. Модели сохранили производительность**:** несмотря на увеличение точности, модели YOLOv7 сохраняют приемлемую скорость обработки, что делает их подходящими для решения задач, требующих высокой точности без значительного увеличения вычислительных затрат. Считается, что модели семейства YOLOv7 идеально подходят для задач, где требуется максимальная точность детектирования, таких как медицинская диагностика, анализ изображений высокого разрешения и сложные сцены с большим количеством объектов.

Постановка задачи исследований

Для проведения комплексных исследований были выбраны, как это показано выше в результате проведенного краткого анализа, современные модели семейств YOLOv5 и YOLOv7. Эффективность этих моделей по точности детектирования объектов на изображениях и по скорости их вычисления необходимо выявить в ходе комплексных исследований. Из этих семейств для исследований выбраны:

- YOLOv5s (одна из наиболее быстродействующих моделей из семейства YOLOv5) и YOLOv5x (наиболее точная модель из этого семейства).
- YOLOv7 (базовая модель семейства YOLOv7) и YOLOv7x (расширенная модель - с большим числом сверточных слоев - и поэтому наиболее точная модель семейства).

Цель комплексных исследований этих четырех моделей – выявить точность детектирования ЛО четырех классов на изображениях и оценить скорость вычисления каждой из этих моделей. По сути, в результате исследований следует определить, какая из четырех моделей наиболее эффективна при решении задач детектирования ЛО на изображениях.

В качестве ЛО на изображениях используются объекты следующих четырех классов:

- Птица (bird).
- Беспилотный летательный аппарат (БПЛА) самолётного типа (aircraft-type_uav).
- БПЛА вертолётного типа (helicopter-type_uav).
- Неизвестный объект (unknown).

Для оценки точности детектирования ЛО на изображениях используются метрики: precision, recall, AP@50, AP@50:95 и для всех классов ЛО - mAP@50, mAP@50:95. Оценка времени обработки и анализа изображений для каждой модели осуществляется по метрике FPS - количество обрабатываемых и анализируемых изображений в секунду.

Разработанный датасет

Для проведения исследования нами был создан набор данных (датасет), включающий размеченные RGB-изображения ЛО упомянутых выше четырех классов. При формировании набора данных использовались изображения из общедоступных датасетов платформы Kaggle [8], а также самостоятельно отобранные изображения. Важно отметить, что объекты

представлены в различных размерах, условно делящихся на малые (площадь не более 32x32 пикселя), средние (от 32x32 до 96x96 пикселей) и крупные (свыше 96x96 пикселей). Датасет состоит из 3430 изображений размером 640x640 пикселей, на которых присутствуют как одиночные ЛО, так и множественные, в том числе разных классов (всего размечено 5880 объектов). При решении задачи изменения дисбаланса числа изображений по классам применялись методы аугментации. Размеченные на изображениях летающие объекты распределились следующим образом: 1870 БПЛА вертолетного типа, 1748 самолетного, 1402 птицы и 860 неизвестных объектов.

Обучение и валидация моделей из семейств YOLOv5 и YOLOv7

Обучение проводилось на онлайн-платформе Kaggle с использованием вычислительного устройства GPU T4 x2. Платформа предоставляет до 30 часов в неделю бесплатно, что позволило провести обучение и комплексные исследования выбранных четырех моделей.

*Параметры обучения***.** Для объективности сравнения все модели обучались с одинаковыми гиперпараметрами:

- Batch size = 32 – оптимальный размер для баланса между скоростью и стабильностью
- Image Size = 640x640 – в соответствии с размером изображений в датасете
- Epochs = 100 – для глубокого анализа изображения (динамика обучения).

*Результаты обучения***.** Процесс обучения проходил одинаково для всех моделей: потери уменьшались с каждой пройденной эпохой, а значения метрик росли. Также на графиках рисунка 1 нельзя заметить ярких признаков переобучения или выхода на плато, что говорит о целесообразности дальнейшего обучения.

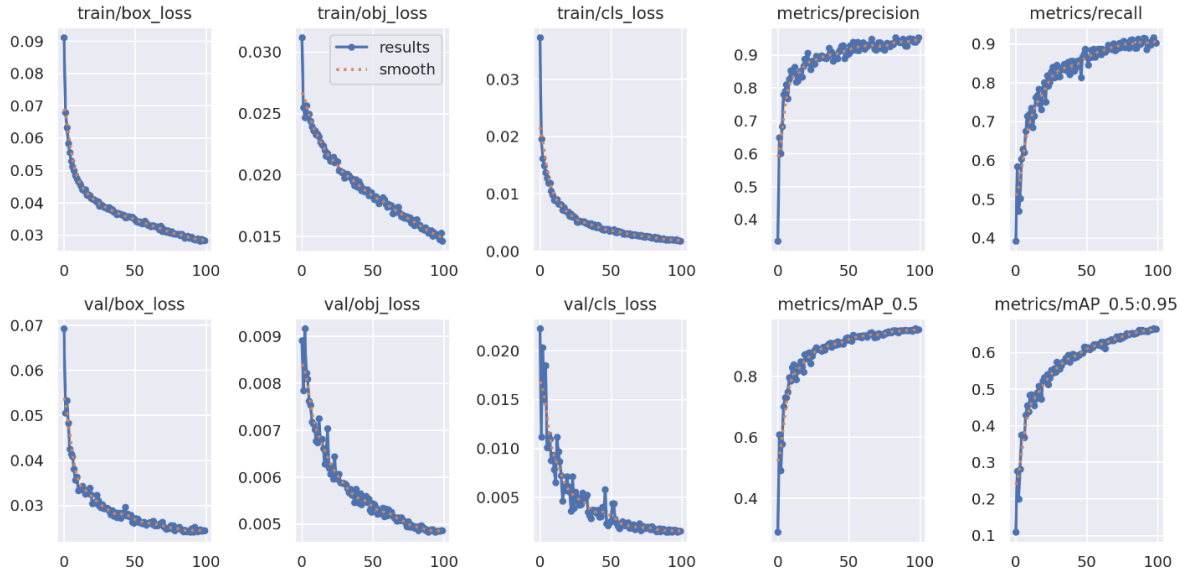

Рис. 1. Графики значений функции потерь и метрик в процессе обучения модели YOLOv5s

Информация о времени обучения каждой модели приведена в таблице 1.

Таблица 1. Сводные данные по времени обучения моделей СНС

Модель	YOLOv5s	YOLOv5x	YOLOv7	YOLOv7x
Время обучения (в часах)	6.97	5.96	18.71	28.37

Неожиданным результатом является превосходство по скорости обучения старшей модели YOLOv5x над младшей YOLOv5s. Возможно это связано с более эффективным разделением нагрузки на два графических ускорителя, в то время как модель YOLOv5s обучалась на одном ускорителе. Также можно видеть, что модели семейства YOLOv7 обучаются значительно дольше моделей младшего семейства, возможно это связано с менее

оптимизированной имплементацией. Так как обучение этих моделей выходило за рамки сессии в 12 часов, использовался параметр resume, чтобы продолжать обучение в новой сессии.

Все модели были провалидированы по итогам обучения. Результаты валидации представлены в виде значений метрик точности детектирования ЛО в таблице 2.

Таблица 2. Результаты по точности детектирования ЛО при валидации моделей

Class and Model / Metric	Precision	Recall	AP@50	AP@50:95
Aircraft YOLOv5s	**0.936**	**0.912**	**0.946**	**0.673**
Aircraft YOLOv5x	**0.978**	**0.929**	**0.981**	**0.777**
Aircraft YOLOv7	**0.935**	**0.834**	**0.931**	**0.712**
Aircraft YOLOv7x	**0.968**	**0.879**	**0.947**	**0.723**
Bird YOLOv5s	0.938	0.941	0.977	0.645
Bird YOLOv5x	0.966	0.974	0.985	0.730
Bird YOLOv7	0.923	0.921	0.917	0.615
Bird YOLOv7x	0.964	0.885	0.934	0.627
Helicopter YOLOv5s	**0.989**	**0.929**	**0.961**	**0.705**
Helicopter YOLOv5x	**0.982**	**0.932**	**0.977**	**0.777**
Helicopter YOLOv7	**0.967**	**0.897**	**0.964**	**0.691**
Helicopter YOLOv7x	**0.978**	**0.920**	**0.966**	**0.708**
Unknown YOLOv5s	0.894	0.884	0.941	0.638
Unknown YOLOv5x	0.957	0.929	0.968	0.760
Unknown YOLOv7	0.959	0.848	0.942	0.674
Unknown YOLOv7x	0.901	0.875	0.939	0.705
Class and Model / Metric	Precision	Recall	mAP@50	mAP@50:95
All YOLOv5s	**0.939**	**0.916**	**0.956**	**0.665**
All YOLOv5x	**0.971**	**0.941**	**0.978**	**0.761**
All YOLOv7	**0.946**	**0.875**	**0.938**	**0.673**
All YOLOv7x	**0.953**	**0.890**	**0.946**	**0.691**

По результатам из этой таблицы можно считать, что при валидации модель YOLOv5x лидирует по всем ключевым метрикам точности детектирования: mAP@50=0.978, mAP@50:95=0.761, Precision=0.971, Recall=0.941.

Результаты исследования моделей СНС

Все модели были исследованы (тестированы) с использованием одинаковой тестовой выборки. Ниже приведена таблица 3, в которой содержится время, затраченное каждой моделью на тестирование 514 изображениями размера 640x640 пикселей, среднее время, затраченное на обработку одного изображения из тестового набора данных, а также количество обрабатываемых изображений в секунду (метрика FPS) каждой модели, рассчитанное на основании среднего времени обработки одного изображения.

Таблица 3. Сводные данные по скорости работы моделей

Модель	Общее время тестирования (сек)	Время, затрачиваемое на одно изображение (мс)	FPS
YOLOv5s	19.26	37.47	26.7
YOLOv5x	40.51	78.81	12.7
YOLOv7	27.61	53.72	18.6
YOLOv7x	35.06	68.21	14.7

Несмотря на преимущество по точности детектирования ЛО модели YOLOv5x над остальными моделями – она способна обрабатывать наименьшее количество кадров в секунду

среди всех исследуемых моделей. Этот недостаток, в зависимости от задач, стоящих перед моделью, может стать весомым доводом в пользу выбора более легковесных моделей. Также важно отметить, что модель YOLOv5s с большим отрывом лидирует по скорости работы, при этом не сильно проигрывая в точности, что делает ее хорошим выбором для задач детектирования и классификации объектов в режиме реального времени.

По итогам тестирования моделей были получены значения точности детектирования ЛО по метрикам Precision, Recall, AP@50 и AP@50:95. Эти значения отражены в таблице 4.

Таблица 4. Результаты по точности детектирования ЛО при тестировании моделей

Class and Model / Metric	Precision	Recall	AP@50	AP@50:95
Aircraft YOLOv5s	**0.980**	**0.848**	**0.941**	**0.676**
Aircraft YOLOv5x	**0.970**	**0.893**	**0.964**	**0.768**
Aircraft YOLOv7	**0.964**	**0.856**	**0.957**	**0.732**
Aircraft YOLOv7x	**0.972**	**0.852**	**0.954**	**0.735**
Bird YOLOv5s	0.974	0.893	0.942	0.676
Bird YOLOv5x	0.988	0.938	0.967	0.729
Bird YOLOv7	0.949	0.875	0.909	0.647
Bird YOLOv7x	0.930	0.883	0.940	0.683
Helicopter YOLOv5s	**0.908**	**0.915**	**0.899**	**0.642**
Helicopter YOLOv5x	**0.921**	**0.932**	**0.918**	**0.719**
Helicopter YOLOv7	**0.930**	**0.894**	**0.910**	**0.668**
Helicopter YOLOv7x	**0.926**	**0.880**	**0.938**	**0.696**
Unknown YOLOv5s	0.924	0.870	0.938	0.677
Unknown YOLOv5x	0.965	0.935	0.979	0.770
Unknown YOLOv7	0.934	0.914	0.945	0.685
Unknown YOLOv7x	0.905	0.870	0.928	0.708
Class and Model / Metric	Precision	Recall	mAP@50	mAP@50:95
All YOLOv5s	**0.947**	**0.882**	**0.930**	**0.668**
All YOLOv5x	**0.961**	**0.925**	**0.957**	**0.747**
All YOLOv7	**0.944**	**0.930**	**0.930**	**0.683**
All YOLOv7x	**0.933**	**0.871**	**0.940**	**0.706**

Для удобства при детальном анализе результатов из таблицы 4 усредненные результаты для всех классов ЛО по точности детектирования по метрике mAP@50 приведены в виде столбчатой диаграммы на рисунке 2.

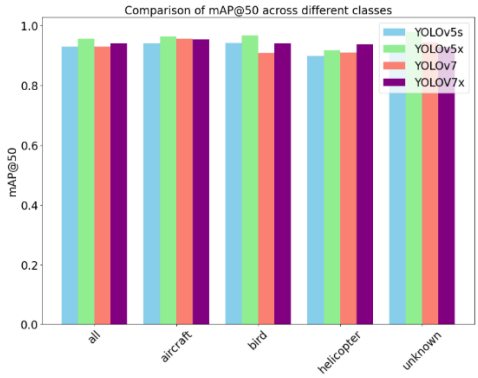

Рис. 2. Столбчатая диаграмма значений метрики mAP@50 для всех классов ЛО и значенийметрики AP@50 *для каждого класса объектов по итогам тестирования всех моделей СНС*

Полученные данные подтверждают сделанный по результатам валидации вывод о превосходстве по точности детектирования модели YOLOv5x над другими исследуемыми

моделями. Однако, не стоит забывать о вычислительной сложности данной модели. В некоторых задачах компьютерного зрения скорость работы модели является приоритетным показателем, что может служить поводом для выбора при решении таких задач более быстрых моделей СНС, таких как, например, YOLOv5s.

Заключение

Наибольшую точность детектирования ЛО на изображениях валидационной выборки датасета продемонстрировала модель YOLOv5x, достигнув следующих значений по метрикам: mAP@50 = 0.978 и mAP@50:95 = 0.761. Аналогичный результат получен и на тестовой выборке. Это подтверждает её эффективность при решении задач, где критически важна точность детектирования. При этом модель YOLOv5s показала наилучшие результаты по скорости обработки изображений - 26.7 кадров в секунду, что делает её предпочтительной при выборе моделей СНС для систем компьютерного зрения реального времени.

Сравнительный анализ результатов исследований моделей семейств YOLOv5 и YOLOv7 выявил, что, несмотря на использование современных архитектурных решений в моделях семейства YOLOv7, эти модели требуют значительно большего времени на обучение и уступают модели YOLOv5x по точности детектирования ЛО. Однако модели семействаYOLOv7 демонстрируют более высокую точность детектирования по сравнению с моделью YOLOv5s, что может быть значимо для решения специализированных задач с менее жесткими требованиями к скорости вычисления моделей СНС.

Полученные результаты свидетельствуют: для большинства практических задач мониторинга воздушного пространства оптимальным решением может являться модель YOLOv5s, обеспечивающая баланс между точностью детектирования ЛО и скоростью ее вычисления. Перспективными направлениями дальнейших исследований могут стать оптимизация всех изучаемых здесь моделей для использования во встраиваемых системах и изучение влияния расширенных методов аугментации на точность детектирования мелких ЛО малых размеров.

Проведенные исследования подтвердили, что модели семейства YOLOv5, в первую очередь, YOLOv5x и YOLOv5s, остаются эффективными инструментами для решения задач детектирования объектов воздушного пространства, а выбор конкретной модели из этого семейства должен определяться требованиями к ее точности при детектированиии ЛО на изображениях и к скорости ее вычислений, а также вычислительными ресурсами проектируемой системы компьютерного зрения в каждом конкретном случае.

Список использованных источников

1. Свёрточная нейронная сеть // Хабрахабр. – [Электронный ресурс]. – URL: habr.com/ru/articles/309508/ (дата обращения 01.04.2025).
2. Как работает Object Tracking на YOLO и DeepSort // Хабрахабр. – [Электронный ресурс]. – URL: habr.com/ru/articles/514450/ (дата обращения 01.04.2025).
3. Подготовка датасета для машинного обучения // Хабрахабр. – [Электронный ресурс]. – URL: habr.com/ru/articles/684580/ (дата обращения 01.04.2025).
4. Ultralytics // YOLOv5: Документация. – [Электронный ресурс]. – URL: docs.ultralytics.com/models/yolov5/ (дата обращения 01.04.2025).
5. Github // YOLOv5 in PyTorch > ONNX > CoreML > TFLite // Репозиторий. – [Электронный ресурс]. – URL: github.com/ultralytics/yolov5 (дата обращения 01.04.2025).
6. Ultralytics // YOLOv7: Документация. – [Электронный ресурс]. – URL: docs.ultralytics.com/models/yolov5 (дата обращения 01.04.2025).
7. Github // Implementation of paper - YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors // Репозиторий. – [Электронный ресурс]. – URL: github.com/WongKinYiu/yolov7 (дата обращения 01.04.2025).
8. Kaggle // Datasets // Сайт. – [Электронный ресурс]. – URL: kaggle.com/datasets (дата обращения 01.04.2025).