

Ли Цзюмин (Китай)

Томский Политехнический университет, г. Томск

Научный руководитель: Мамонова Т.Е., канд. техн. наук, доцент

ДЕТЕКТИРОВАНИЕ ОБЪЕКТОВ С ИСПОЛЬЗОВАНИЕМ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ YOLO11

Аннотация

В данной работе выполнено детектирование транспортных средств на аэрофотоснимках с использованием модели сверточной нейронной сети YOLO11. Экспериментальные исследования проведены с использованием наборов данных CARPK, в котором было реализовано детектирование и подсчет количества автомобильных объектов на изображении. Результаты показывают, что модель YOLO11 обладает способностью точно обнаруживать и определять местоположение автомобильных целей на сложных видах аэрофотосъемки.

Ключевые слова: детектирование объектов, сверточная нейронная сеть, YOLO11, CARPK

Введение

В настоящее время детектирование объектов на базе аэрофотоснимков имеет широкие перспективы применения в области контроля и безопасности транспортного потока. Задача детектирования объектов является важной частью применения технических устройств, на базе которых в качестве платформы в сочетании с технологией компьютерного зрения можно быстро воспринимать окружающую обстановку и сократить потребление человеческого ресурса. В связи с ограничением аппаратных ресурсов оборудования подобных технических устройств, для повышения скорости детектирования и обеспечения его точности применяются облегченные методы глубокого обучения, представленные сверточными нейронными сетями (CNN).

Методы детектирования объектов на основе CNN делятся на две группы:

1. Двухэтапные алгоритмы детектирования на основе генерации регионов, представленные R-CNN и Faster-RCNN.
2. Одноэтапные алгоритмы детектирования на основе регрессии, представленные серией YOLO (you only look once) [1].

Одноэтапные методы детектирования имеют более высокую скорость определения объекта и сквозную производительность по сравнению с двухэтапными методами. В связи с этим в настоящее время данные методы используются в качестве основного объекта исследования при детектировании объектов с применением результатов аэрофотосъемки.

Алгоритм серии YOLO с момента своего появления в 2015 году занял важное место в области детектирования объектов в реальном времени благодаря

эффективной скорости и высокой точности. YOLO11 является последней версией серии, которая улучшает способность модели к извлечению признаков и повышает точность детектирования за счет внедрения архитектурных оптимизаций, таких как новые модули C3k2 и C2PSA, сохраняя при этом легкие характеристики и оптимизированную скорость вывода.

Таким образом, в представляемой работе в качестве объекта исследования используются структура модели сверточной нейронной сети YOLO11 и представленные значения весовых коэффициентов нейронной сети YOLO11n. Обучение модели выполняется на основе набора данных CARPK для детектирования объектов в сложных сценариях на малых высотах. В качестве инициализирующих параметров используются предварительно обученные весовые коэффициенты модели YOLO11n.

Описание YOLO11 model

Модель YOLO11 является последней версией семейства YOLO, разработанной командой Ultralytics в США и Испании. Модель YOLO11 отличается от предыдущих версий YOLO, в том числе измененной архитектурой модели, улучшением извлечения признаков и оптимизацией методов обучения. По сравнению с моделью YOLOv8, в YOLO11 модуль CF2 заменен на C3K2, что позволяет использовать различные размеры сверточных ядер и адаптироваться к более сложным задачам и требованиям многомасштабного обнаружения. За модулем SPPF был добавлен модуль C2PSA, целью которого является улучшение извлечения признаков за счет добавления механизма многоголового внимания. Архитектура модели YOLO11 приведена на рисунке 1.

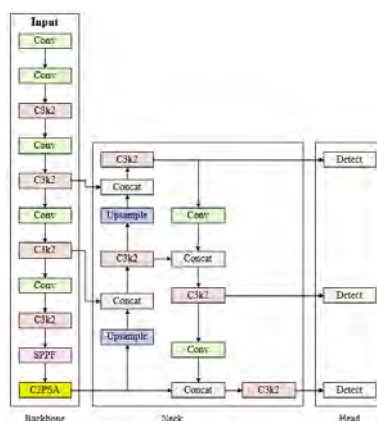


Рис. 1. Архитектура нейронной сети YOLO11

Для обеспечения гибкого применения на аппаратных устройствах в различных сценариях, модель YOLO11 разделена на пять моделей с различными масштабами в соответствии с параметрами ширины и глубины: YOLO11n, YOLO11s, YOLO11m, YOLO11l и YOLO11x. Согласно данным из официальной документации Ultralytics, указанной в [2], соответствующие глубина канала

Depth, ширина *Width*, максимальное количество каналов *Max channels*, количество параметров *Parameters*, вычислительная сложность *GFLOPs* и средняя точность *mAP* для этих пяти моделей приведены в таблице 2. Как видно из данных таблицы, с увеличением размера модели растет количество параметров (от 2,6М до 56,9М), увеличивается вычислительная сложность (от 6,5 до 194,9 GFLOPs), а среднее значение средней точности, рассчитанное при различных пороговых значениях IoU, варьирующихся от 0,50 до 0,95 (*mAP@0.5:0.95*) повышается с 39,5 до 54,7 %.

Таблица 1

Параметры YOLO11 для различных размеров объектов

Модель	Depth	Width	Max channels	Parameters(M)	GFLOPs	mAP@0.5:0.95, %
YOLO11n	0,50	0,25	1024	2,6	6,5	39,5
YOLO11s	0,50	0,50	1024	9,4	21,5	47,0
YOLO11m	0,50	1,00	512	20,1	68,0	51,5
YOLO11l	1,00	1,00	512	25,3	86,9	53,4
YOLO11x	1,00	1,50	512	56,9	194,9	54,7

Из-за наличия разноплановых параметров модели, высоких требований к вычислениям и хранению данных, высокой потребности в памяти компьютера и необходимости ускорения с помощью высокопроизводительной видеокарты для использования всех преимуществ моделей с большим количеством параметров, например, YOLO11m, YOLO11l и YOLO11x, их трудно развернуть на устройствах с ограниченной производительностью и вычислительными ресурсами. Исходя из аппаратных ограничений и прикладных требований к платформам, разворачиваемых на технических устройствах, было определено, что модели сверточных нейронных сетей YOLO1n и YOLO1ls являются наиболее подходящими для решения задач распознавания объектов. При этом конкретный выбор требует компромисса между точностью (*mAP*) и скоростью (*Parameters* и *GFLOPs*) в зависимости от сложности сценария.

Эксперимент и результат

Для тестирования работоспособности модели YOLO11n при обнаружении объектов на аэрофотоснимках используется набор данных CARPK (Car Parking Lot Dataset), представленный в источнике [3]. Указанный набор является крупномасштабным набором данных для обнаружения и подсчета автомобилей, он содержит 1 448 изображений и информацию о 89 777 автомобилях в различных сценах для 4 различных парковок. Изображения были получены с высоты 40 м. Для обучения набор данных делится на обучающий, валидационный и тестовый набор по соотношению 7:1:2. Количество изображений, соответствующих каждому набору данных, приведено в таблице 2.

Таблица 2

Расположение набора данных

набор данных	обучающий	валидационный	тестовый	Итого
CARPK	1031	145	289	1448

В таблице 3 показаны ключевые параметры для обучения модели. Для обучения нейронной сети были выбраны следующие параметры: 100 циклов обучения с оптимизатором (*optimizer*) SGD, импульсом (*momentum*) 0,937 и начальной скорости обучения (*lr0*) 0,01. Учитывая, что в процессе обучения используется центральный процессор, для уменьшения объёма занимаемой памяти и времени обучения, размер входных изображений (*imgsz*) установлен на 640×640, а количество изображений в каждой партии (*batchsize*) равно 8.

Таблица 3

Параметры для обучения модели

Название параметра	epochs	batch	imgsz	optimizer	momentum	lr0
Значение параметра	100	8	640	SGD	0.937	0.01

Для проверки производительности модели используются такие показатели оценки, как точность (P), отзыв (R), средняя точность AP и среднее значение средней точности (mAP). Формулы приведены таблице 4.

Таблица 4

Формулы для оценки показателей

Показатель оценивания	Формула	Описание
Точность(P)	$P = TP / (TP + FP)$, где TP - истинно положительный прогноз, FP - ложно положительный прогноз	Точность определяет долю истинно положительных прогнозов среди всех положительных прогнозов, оценивая способность модели избегать ложных срабатываний. Используется для оценки точности предсказаний модели.
Отзыв(R)	$R = TP / (TP + FN)$, где FN – ложно отрицательный прогноз	Отзыв рассчитывает долю истинных положительных результатов среди всех реальных положительных результатов, оценивая способность модели обнаруживать все экземпляры класса. Используется для оценки способности модели находить полную цель.

Средняя точность(AP)	$AP = \int (P(R) dR)$	Средняя точность представляет собой площадь под кривой точность-отзыв и используется для измерения общей эффективности модели при различных пороговых значениях.
Среднее значение средней точности (mAP)	$mAP = \sum(AP_i) / N$ где N – количество категорий	Среднее значение AP всех категорий. Используется для комплексной оценки эффективности модели в сценах обнаружения объектов с несколькими классами.

Таблица 5

Показатели обучения модели

категория	P, %	R, %	mAP0.5, %	mAP0.5-0.95, %	Parameters(M)	GFLOPs
Car	99,10	98,50	99,40	84.00	2,58	6,30

Показатели оценки модели, полученные в конце обучения модели, приведены в таблице 5. Видно, что среднее значение средней точности при пороговых значениях IoU в диапазоне от 0,5 до 0,95 составляет 84 %, а количество параметров только 2,58 миллионов. Это означает, что модель обладает высокой точностью и скоростью обнаружения.

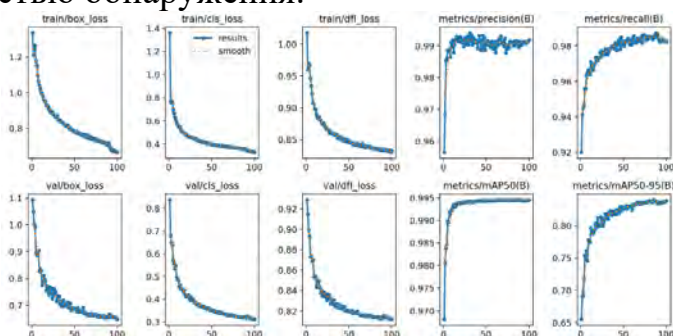


Рис. 2. Показатели оценки модели

На рисунке 2 показаны графики изменений показателей модели YOLO11n во время обучения и валидации. Показатель *train/box_loss* представляет собой изменение потерь ограничивающий прямоугольник (*bouding box*) в процессе обучения, видно, что потери постепенно уменьшаются с увеличением количества тренировок, что указывает на постоянную оптимизацию модели. Показатель *train/cls_loss* отражает точность предсказания модели по целевой категории, потери постепенно уменьшаются в процессе обучения, что указывает на улучшение способности модели определять целевую категорию. Показатели *val/cls_loss* и *val/dfl_loss* отражают улучшение способности модели распознавать категории и находить местоположение цели на валидационном множестве.

Кривые $metrics/mAP50(B)$ и $metrics/mAP50(B)$ показывают, что модель сохраняет лучшую эффективность обнаружения при различных пороговых значениях.

Далее было выполнено тестирование обученной модели на тестовом наборе и подсчет количества автомобилей на каждом изображении, результаты показаны на рисунке 3. Из полученных результатов обнаружения выбираются четыре изображения с различными сценами, как показано на рисунке 4. Видно, что все автомобили на изображении обнаружены правильно.

	A	B	C	D
1	image_name	vehicle_count	detection_path	
2	20160331_NTU_00001.jpg	4	C:\Users\lijiuming\D	
3	20160331_NTU_00003.jpg	4	C:\Users\lijiuming\D	
4	20160331_NTU_00004.jpg	4	C:\Users\lijiuming\D	
5	20160331_NTU_00005.jpg	3	C:\Users\lijiuming\D	
6	20160331_NTU_00007.jpg	3	C:\Users\lijiuming\D	
7	20160331_NTU_00026.jpg	37	C:\Users\lijiuming\D	
8	20160331_NTU_00028.jpg	42	C:\Users\lijiuming\D	
9	20160331_NTU_00029.jpg	42	C:\Users\lijiuming\D	
10	20160331_NTU_00030.jpg	28	C:\Users\lijiuming\D	
11	20160331_NTU_00031.jpg	20	C:\Users\lijiuming\D	
12	20160331_NTU_00034.jpg	12	C:\Users\lijiuming\D	
13	20160331_NTU_00045.jpg	28	C:\Users\lijiuming\D	
14	20160331_NTU_00047.jpg	32	C:\Users\lijiuming\D	
15	20160331_NTU_00055.jpg	43	C:\Users\lijiuming\D	

Рис. 3. Результаты тестирования модели YOLO11n:
а) обнаружения на тестовом наборе, б) подсчет автомобилей

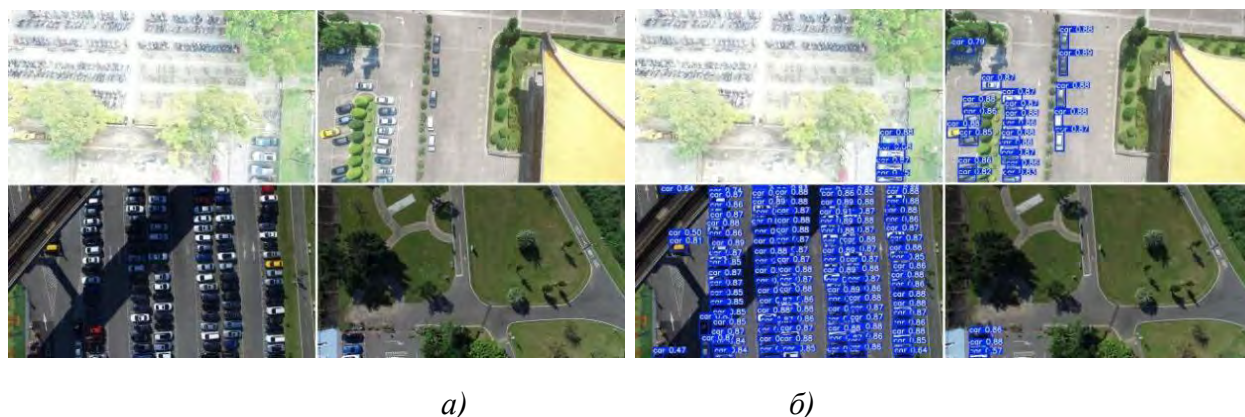


Рис. 4. Результаты визуализации: а) исходное изображение б) обнаруженное изображение

Заключение

В данной работе реализовано обнаружение транспортных средств на аэрофотоснимках, выполненных на расстоянии 40 м., с применением модели сверточной нейронной сети YOLO11n. Результат оценки и визуализации показывают, что модель YOLO11 обладает достаточной производительностью в задаче обнаружения транспортных целей. Полученный результат подтверждает эффективность нейронной сети YOLO11 для быстрого и точного обнаружения и подсчета транспортных средств на сложных аэрофотоснимках, подтверждает осуществимость и преимущества этой технологии в мониторинге транспортно-

го потока в реальном времени, а также обеспечивает сильную техническую поддержку для интеллектуальных систем мониторинга дорожного движения с помощью технических устройств и аппаратных средств, работающих на низких расстояниях от земли.

Список литературы

1. Redmon J., Divvala S., Girshick R., and Farhadi A., «You only look once: Unified, real-time object detection» in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 779–788.
2. Ultralytics «Object Detection Models - YOLO11 Documentation». – URL: <https://docs.ultralytics.com/tasks/detect/#models/> (дата обращения: 05.03.2025)
3. Car Parking Lot Dataset (CARPK). – URL: <https://lafi.github.io/LPN/>

Ли Шухуэй (Китай)

Томский политехнический университет, г. Томск

Научный руководитель: Ремнев Геннадий Ефимович, д.т.н., профессор

ОПТИМИЗАЦИЯ КОНСТРУКЦИИ МИШЕНИ ДЛЯ ГЕНЕРАЦИИ ИМПУЛЬСНЫХ НЕЙТРОННЫХ ПОТОКОВ

Аннотация: В статье рассматривается проблема оптимизации конструкции мишени для генерации импульсных нейтронных потоков, используемых в бор-нейтронозахватной терапии. Определены оптимальные толщины литиевой, фторид-литиевой и оксид-литиевой мишеней для импульсных протонных пучков 2,5 МэВ. Сделан вывод о том, что использование импульсных пучков требует учета их энергетического спектра. Полученные результаты важны для повышения эффективности генерации нейтронов в импульсных ускорительных системах для развития бор-нейтронозахватной терапии.

Ключевые слова: бор-нейтронозахватная терапия, импульсный пучок, мишень, нейтрон, моделирование

Введение

Нейтрон-захватная терапия и ее разновидность – бор-нейтронзахватная терапия (БНЗТ) – это технология таргетной лучевой терапии, реализуемая на клеточном уровне. Механизм БНЗТ заключается во введении целевых молекулярных препаратов, несущих стабильный изотоп бор-10 (^{10}B), специфически удерживающийся в опухолевых клетках, и облучении области опухоли потоком эпитепловых нейтронов, вызывающих ядерную реакцию с высвобождением ус-