

ВЛИЯНИЕ ГИПЕРПАРАМЕТРОВ НЕЙРОННОЙ СЕТИ НА КАЧЕСТВО РАСПОЗНАВАНИЯ ОБЪЕКТОВ В РЕАЛЬНОМ ВРЕМЕНИ

А.А. Рогозин, В.В. Шатских

Межвидовой центр подготовки и боевого применения войск РЭБ
(учебный и испытательный), Тамбов

Аннотация. В данной статье рассматриваются основные этапы обучения модели нейронной сети для определения объектов в реальном времени, а также влияние некоторых гиперпараметров на точность и скорость обучаемой модели на основе архитектуры YOLOv8
Ключевые слова: нейронная сеть, сверточная нейронная сеть, yolov8.

THE EFFECT OF NEURAL NETWORK HYPERPARAMETERS ON THE QUALITY OF OBJECT RECOGNITION

A.A. Rogozin, V.V. Shatskikh

Interspecific Center for Training and Combat Use of Electronic Warfare Troops
(training and testing) Tambov

Abstract.. This article discusses the main stages of training a neural network model for determining objects in real time, as well as the influence of some hyperparameters on the accuracy and speed of the model being trained based on the YOLOv8 architecture
Keywords: neural network, convolutional neural network, yolov8.

Искусственный интеллект (ИИ - artificial intelligence, AI) - это область информатики, в рамках которой разрабатываются компьютерные программы для решения задач, требующих имитации мыслительной деятельности человека. Такие программы могут делать обобщения и выводы, выявлять взаимосвязи и обучаться с учетом накопленного опыта. Системы искусственного интеллекта ни в коем случае не заменяют человека, они расширяют и дополняют его возможности.

Два ключевых понятия в системах искусственного интеллекта - это нейронные сети и машинное обучение. Нейронная сеть по своей сути представляет некую математическую модель и ее программную реализацию, которая в упрощенном виде воссоздает работу биологической нейронной сети человека. А машинное обучение - это набор специальных алгоритмов, благодаря которым воплощается ключевое свойство нейронных сетей - способность самообучения на основе получаемых данных. Чем больший объем информации будет представлен в качестве тренировочного массива данных, тем проще обучающим ал-

горитмам найти закономерности и тем точнее будет выдаваемый результат. Нейронная сеть по своей сути представляет некую математическую модель и ее программную реализацию, которая в упрощенном виде воссоздает работу биологической нейронной сети человека [1].

Пример простейшей нейронной сети: перцептрон. Математически перцептрон можно выразить так (рис. 1)

$$Y = f(x_1 w_1 + x_2 w_2 + b)$$

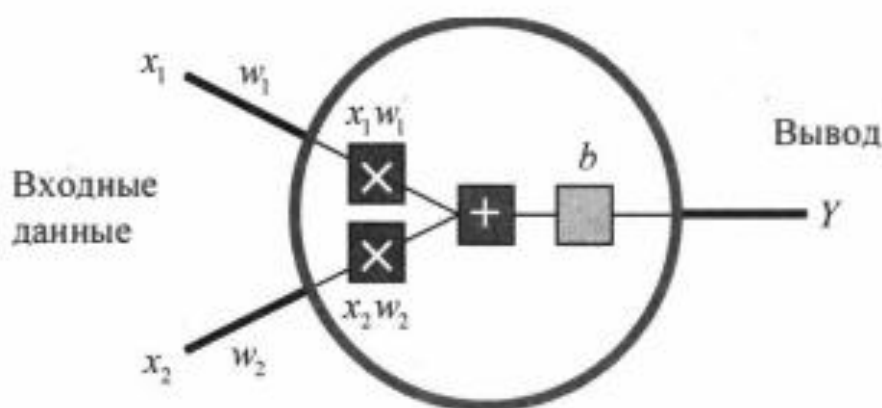


Рисунок 1 – Пример перцептрона

Нейронная сеть дает возможность компьютеру или роботу, мыслить и принимать решения разумно, подобно тому, как думают и действуют люди. Нейронная сеть работает, как и мозг человека, она учится, набирается опыта, а затем на практике использует результаты своего обучения.

Люди давно стремились заставить вычислительные машины мыслить и вести себя так же, как человек, и таким образом научить их решать не свойственные компьютерам задачи — например, играть в шахматы, сочинять стихи и писать музыку.

YOLO – это алгоритм обнаружения объектов в реальном времени, который использует глубокое обучение для распознавания объектов на изображениях и видео. В отличие от других алгоритмов обнаружения объектов, которые выполняют обнаружение в несколько этапов, YOLO выполняет обнаружение в один этап, что делает его намного быстрее, чем другие методы. Алгоритм YOLO делит входное изображение на сетку ячеек и предсказывает ограничивающую рамку и вероятности классов для каждой ячейки. Координаты ограничивающей рамки относятся к координатам ячейки, что означает, что один объект может быть обнаружен в нескольких ячейках.

Ввод: YOLO берет входное изображение или видео и изменяет его размер до фиксированного. Модель глубокого обучения.: YOLO использует глубокую сверточную нейронную сеть (CNN) в качестве своей базовой архитектуры. Сеть обучается на большом наборе данных помеченных изображений, чтобы научиться распознавать объекты. Прогнозирование ограничивающей рамки: Алгоритм YOLO предсказывает ограничивающие рамки для каждого объекта на

входном изображении. Каждая ограничивающая рамка представлена пятью значениями: x , y , w , h и показателем достоверности.

Координаты (x , y) представляют центр прямоугольника, в то время как w и h представляют ширину и высоту прямоугольника соответственно. Показатель достоверности показывает, насколько алгоритм уверен в том, что в поле находится объект (рис. 2) [2].

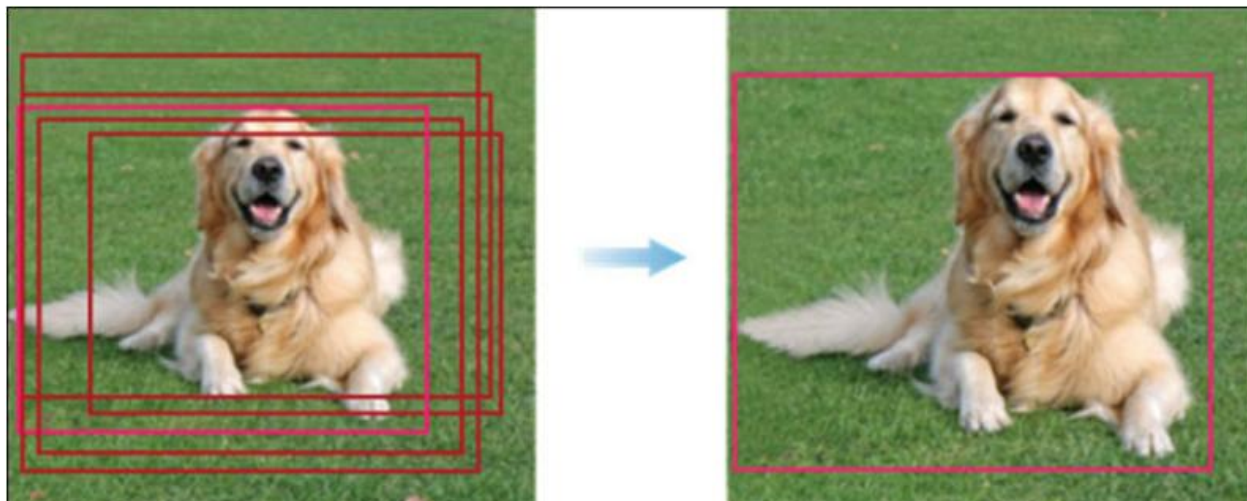


Рисунок 2 – Принцип работы YOLO

Большинство современных детекторов объектов реализуют комбинацию классификации и регрессии ограничивающей рамки. Классификация пытается предсказать класс объекта в области изображения, а регрессия ограничивающей рамки пытается определить область расположения объекта, предсказывая самую узкую рамку, содержащую объект. Рассмотрим эталон ограничительной рамки g , связанный с меткого класса y , и гипотезу обнаружения x ограничительной рамки b . Поскольку b обычно включает в себя объект и некоторое количество фона, может быть трудно определить, правильно ли обнаружен объект. Обычно это решается с помощью метрики пересечения по объединению (intersection over union, IoU).

Если IoU не превышает порога для любого объекта, x считается фоном и обозначается как «отрицательный» пример. Хотя нет необходимости определять положительные/отрицательные примеры для задачи регрессии ограничивающей рамки, порог IoU и также требуется для выбора набора образцов используемого для обучения регрессора. Хотя пороговые значения IoU, применяемые для двух задач, не обязательно должны быть идентичными, на практике это обычное дело. Следовательно, порог IoU и определяет качество детектора. Большие пороги способствуют тому, чтобы обнаруженные ограничивающие рамки были точно совмещены с эталонами. Небольшие пороги вознаграждают детекторы, которые создают свободные ограничивающие рамки с небольшим перекрытием с эталоном. Некоторые примеры гипотез повышения качества показаны на рис. 3 [3].

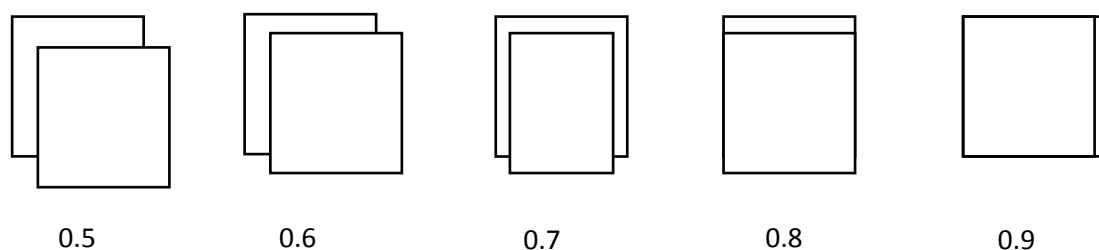


Рисунок 3 – Примеры повышения качества. Числа представляют собой значения IoU между двумя ограничивающими прямоугольниками, указывая, насколько хорошо они совмещены друг с другом

Одним из ключевых способов улучшить качество модели YOLOv8 является оптимизация гиперпараметров. Например, можно настроить параметры обучения, такие как скорость обучения, размер пакета и оптимизатор чтобы достичь лучших результатов. Также можно выбрать другую версию архитектуры YOLOv8, она отличается скоростью обработки изображения и точностью. Результаты влияния различных гиперпараметров на точность модели указано на рис. 4.

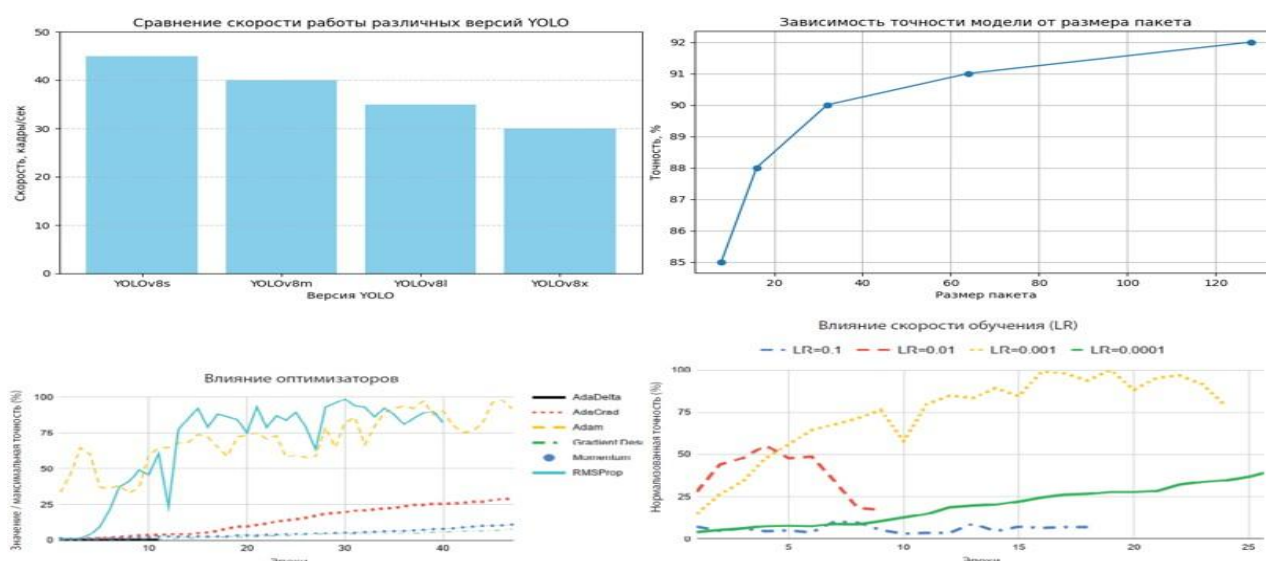


Рисунок 4 – Результаты влияния различных гиперпараметров на точность обучаемой модели

Список литературы

1. Компьютерное зрение. Современные методы и перспективы развития / ред. Р. Дэвис, М. Терк; пер. с англ. В. С. Яценкова. – М.: ДМК Пресс, 2022. – 690 с.: ил.
2. Искусственный интеллект и компьютерное зрение. Реальные проекты на Python, Keras и TensorFlow. — СПб.: Питер, 2023. — 624 с.: ил. — (Серия «Бестселлеры O'Reilly»).
3. Основы искусственного интеллекта в примерах на Python. Самоучитель. — 2-е изд., перераб. и доп. — СПб.: БХВ-Петербург, 2024. — 448 с.: ил. — (Самоучитель).

References

1. Computer vision. Modern methods and prospects of development / ed. R. Davis, M. Turk; translated from English by V. S. Yatsenkova. – M.: DMK Press, 2022. – 690 p.: ill.
2. Artificial intelligence and computer vision. Real projects in Python, Keras and TensorFlow. — St. Petersburg: St. Petersburg, 2023. — 624 p.: ill. — (Series "O'Reilly Bestsellers").
3. The basics of artificial intelligence in Python examples. Self-instruction. — 2nd ed., reprint. and additional — St. Petersburg: BHV-Petersburg, 2024. — 448 p.: ill. — (Tutorial)