

НЕЙРОСЕТЕВОЙ АЛГОРИТМ ОБНАРУЖЕНИЯ И КЛАССИФИКАЦИИ ОБЪЕКТОВ НА ИЗОБРАЖЕНИИ

студент Ушаков П. А.

Научный руководитель - канд. физ-мат. наук Козлова Е. И.

Белорусский государственный университет

Минск, Беларусь

Введение

Разработка систем, анализирующих видеопоток в режиме реального времени, представляет собой актуальную и активно развивающуюся теоретическую и практическую проблему во многих областях. Актуальность данной работы заключается в том, что нейронные сети могут применяться в различных интерактивных системах, умных городах, видеонаблюдении и многих других областях. Кроме того, их использование позволяет автоматизировать процесс и сократить время на выполнение задач.

Для обнаружения и классификации объектов на изображении могут применяться различные типы нейронных сетей:

Сверточные нейронные сети (CNN): наиболее часто используемый тип сетей в обработке изображений, поскольку они автоматически и эффективно извлекают признаки из изображений на разных уровнях, что критически важно для точного распознавания.

R-CNN (Region-based Convolutional Neural Networks) и их вариации (Fast R-CNN, Faster R-CNN): эти сети обеспечивают более обнаружение объектов, путем сочетания CNN для извлечения признаков и регионов интереса на изображениях.

SSD (Single Shot MultiBox Detector): эти сети также обрабатывают изображение за один проход и могут обнаруживать объекты разных масштабов.

Сбор данных и формирование датасета

Сбор информации для датасета может проводиться различными

способами в зависимости от задачи и типа требуемых данных. Для многих задач уже есть готовые датасеты, которые можно найти в открытом доступе на различных веб-ресурсах.

Формирование качественного датасета является одной из самых важных задач в построении модели обнаружения и классификации объектов. Качество и точность определения объектов в поставленной задаче напрямую зависит от данных, которые находятся в датасете. Чем лучше он будет подготовлен, тем меньше времени потребуются на отладку модели, ее тренировки, поиск и устранение неточностей распознавания.

Некоторые задачи требуют ручного ввода данных от пользователей. В таком случае предоставляются разработанные веб-формы и системы, позволяющие проводить аннотацию(разметку) данных. В данной работе использовалась специальная платформа Roboflow. Платформа упрощает процесс разработки моделей для компьютерного зрения, имеет широкую совместимость и гибкость развертывания, у нее простая настройка и интеграция, а также одно из самых значительных ее преимуществ – это безопасность. Пример размеченного датасета представлен на рисунке 1.



Рисунок 1. Размеченный датасет

Методы обнаружения

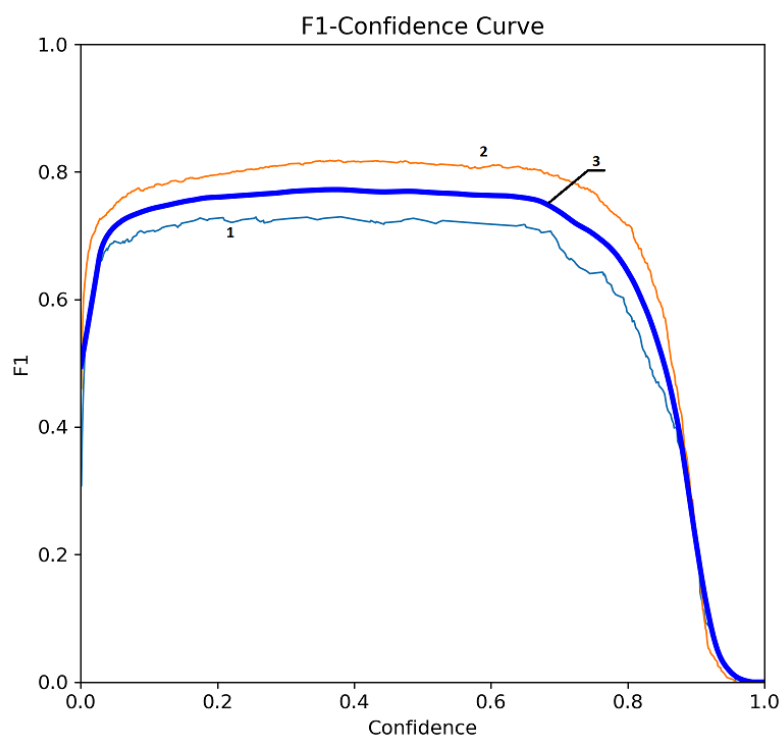
Методы обнаружения объектов, как правило, основаны на машинном либо глубоком обучении. Для методов, основанных на глубоком обучении, свойственно использование сверточных нейронных сетей. В соответствии с

этим, для данной работы был выбран метод YOLO, поскольку к его достоинствам можно отнести скорость, высокую точность, универсальность.

Результаты

Один из наиболее важных критериев оценки работы нейронной сети в случае распознавания изображений – это качество распознавания изображений. Для оценки качества распознавания изображения при помощи нейронной сети мы можем использовать разные метрики. Проверка работоспособности предлагаемого в данной работе алгоритма была исследована на датасете, который содержал изображения максимально приближенные к реальным условиям. Задачей модели обнаружения в алгоритме было нахождение на видео интересующих объектов – человек и сумка/рюкзак.

Исходя из данных, полученных на графике F1-Confidence Curve (Рисунок 2), можно сделать вывод, что максимальное значение F1 для обоих классов достигается при определённом уровне уверенности и составляет примерно 0.77 при пороге уверенности 0.37. Это указывает на то, что модель достигает наилучшего баланса между точностью и полнотой при данном значении confidence. После этой точки F1-мера падает, что может свидетельствовать о уменьшении полноты при более высоких уровнях уверенности.



*1 – объект класса bag, 2 – объект класса person,
3 – для двух классов*

Рисунок 2. F1-Confidence Curve

Также для наглядности работы модели была построена матрица ошибок (Рисунок 3).

Согласно данной матрице, можно сделать вывод, что модель достаточно хорошо справляется с обнаружением человека на изображении, точность предсказания составила 83%, но 17% были ошибочно приняты за класс "background". Для класса "bag", модель предсказала правильно в 67% случаев, однако 32% были ошибочно классифицированы как "background".

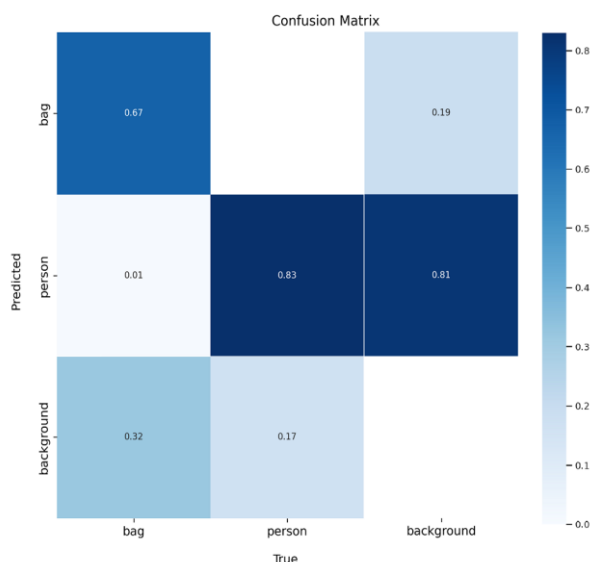


Рисунок 3. Confusion Matrix

Заключение

Распознавание объектов по-прежнему остается одной из самых важных сфер применения для глубокого обучения и компьютерного зрения на сегодняшний день. В работе были подробно рассмотрены основные этапы процесса применения нейронных сетей для классификации объектов, включая подготовку данных. Важным этапом работы с нейронными сетями является выбор подходящих метрик оценки качества классификации. Результаты, полученные с помощью этих метрик, подчеркнули сильные стороны модели в определенных аспектах обнаружения объектов, а также выявили области, требующие улучшения.

Литература

1. Ultralytics YOLOv8 Docs [Электронный ресурс] – Режим доступа: <https://docs.ultralytics.com/> – Дата доступа: 15.12.2023
2. Performance Metrics in Machine Learning [Электронный ресурс] – Режим доступа: <https://neptune.ai/blog/performance-metrics-in-machine-learning-complete-guide> - Дата доступа: 15.12.2023
3. Задача нахождения объектов на изображении [Электронный

ресурс] – Режим доступа:

https://neerc.ifmo.ru/wiki/index.php?title=Задача_нахождения_объектов_на_изображении – Дата доступа: 14.12.2023.