

ИСПОЛЬЗОВАНИЕ НЕЙРОННОЙ СЕТИ MASK R-CNN ДЛЯ РАСПОЗНАВАНИЯ ОБЪЕКТОВ

Е. В. Антонович

Белорусский государственный университет, г. Минск;

antonovich.egor1@gmail.com

науч. рук. – А. М. Сорока, старший преподаватель

В данной статье рассматривается способ использования свёрточной нейронной сети Mask R-CNN на базе модели DETECTRON (разработана подразделением Facebook AI Research) для распознавания объектов. Обучение проводилось на датасете (база данных) из изображений, которые содержат классы объектов, влияющих на дорожную обстановку. Статья содержит описание этапов обучения нейронной сети и тестирования обученной модели. Также предоставлена сводная таблица результатов тестирования обученной модели, содержащая точность и количество ложных срабатываний нейронной сети по каждому классу. Произведено тестирование модели до и после обучения для сравнительного анализа обучаемости нейронной сети на новые классы объектов без ухудшения точности для уже имеющихся классов. Результаты сравнения моделей вынесены в сводную таблицу.

Ключевые слова: искусственные нейронные сети; Mask R-CNN; Detectron; компьютерное зрение; обучение нейронной сети.

ВВЕДЕНИЕ

В наше время большую популярность и широкое применение всё больше и больше получают искусственные нейронные сети (ИНС). Модели нейронных сетей применяются практически во всех сферах жизнедеятельности человека.

Такая область искусственного интеллекта (ИИ), как компьютерное зрение (КЗ), получила своё применение достаточно недавно. Особенно широкое применение КЗ получило в автомобилях, оснащённых автопилотом, для анализа дорожной обстановки путём распознавания объектов дорожного движения с использованием ИНС для обучения высокоуровневых моделей.

Использование нейронных сетей в автопилотах во многом превосходит использование различных лидаров, радаров и датчиков. Как пример можно привести то, что при применении нейросетевых технологий нет ограничения в расстоянии до объектов. Как следствие, это может увеличить время для принятия решений при опасных ситуациях на дороге.

ЦЕЛЬ ИССЛЕДОВАНИЯ

Целью данной научно-технической работы было изучить возможность применения свёрточной нейронной сети (НС) Mask R-CNN в це-

лях анализа дорожной обстановки для автомобилей, оснащённых автопилотами.

Для достижения данной цели были выполнены следующие задачи:

- Собран датасет (база данных (БД)) изображений, содержащий классы объектов, которые способны влиять на дорожную обстановку и поведение автомобиля с автопилотом на дороге.
- Обучена свёрточная нейронная сеть Mask R-CNN на базе модели DETECTRON (разработана подразделением Facebook AI Research).
- Произведено тестирование обученной модели для определения точности распознавания объектов, влияющих на дорожную обстановку.

ОБУЧЕНИЕ И ТЕСТИРОВАНИЕ

Для обучения НС было использовано две БД: COCO (Common Objects in Context) и база данных, которая была собрана из внешних источников, содержащая классы объектов, которые способны влиять на состояние дорожной обстановки.

Собранная база данных была размечена на 24 класса объектов. Были взяты 17 классов из датасета COCO и 7 новых классов.

Список классов, присутствующих в собранном датасете:

- COCO классы: автомобиль, автобус, грузовик, поезд, пешеход, велосипед, мотоцикл, знак «СТОП», светофор, лавочка, пожарный кран, кот, собака, зонт, чемодан, рюкзак, ручная кладь.
- Добавленные классы: автобусная остановка, дорожные работы, конус безопасности, пешеходный переход, велосипедная полоса, дорожный перекрёсток, предупреждение о дорожных работах.

Для добавления новых классов были переинициализированы некоторые слои НС Mask R-CNN.

Общее количество данных для обучения нейронной сети составило:

1. 10050 изображений;
2. 92940 объектов.

ОСУЩЕСТВЛЕНИЕ ТЕСТИРОВАНИЯ ОБУЧЕННОЙ МОДЕЛИ

Тестирование обученной модели производилось на разработанном бенчмарке (система для анализа точности распознавания нейронной сети), написанном на языке программирования Python. Были проведены тесты точности распознавания добавленных классов и классов пешеход, автомобиль, велосипед. Помимо точности распознавания в процентном соотношении было вычислено количество ложных срабатываний НС на

100 кадров. В ходе тестирования был выбран confidence, равный 0,7, значение IoU, равное 0,5 (50%). Результаты приведены в табл. 1.

Таблица 1

Результаты тестирования обученной модели

Название класса	Точность, %	Ложные срабатывания
Автобусная остановка	81	0.8
Дорожные работы	61	5
Конус безопасности	74	12
Пешеходный переход	55	8
Велосипедная полоса	41	13
Дорожный перекрёсток	44	18
Предупреждение о дорожных работах	31	1
Пешеходы	66	42
Автомобили	67	72
Велосипеды	62	4

СРАВНЕНИЕ МОДЕЛЕЙ

Помимо проведённых тестов на классах обученной модели было произведено тестирование предобученной модели на классах автомобиль, пешеход, велосипед. Далее были сравнены результаты до и после обучения. Данные о тестировании приведены в сравнительной таблице 2.

Таблица 2

Сравнение моделей до и после обучения

Название классов	Модель до обучения		Модель после обучения	
	Точность, %	Ложные срабатывания	Точность, %	Ложные срабатывания
Автомобили	42	2	60	34
Пешеходы	59	13	60	20
Велосипеды	35	0,2	56	2,6

ЗАКЛЮЧЕНИЕ

Полученные результаты свидетельствуют о возможности добавления новых классов в модель Detectron и дообучении HC Mask R-CNN.

На базе полученных результатов возможно построение более высокоуровневых моделей, применение в исследовательских целях, а так же

для создания систем анализа дорожной обстановки для автомобилей с автопилотами.

Низкая точность распознавания таких классов, как «перекрёсток», «велосипедная полоса», «предупреждение о дорожных работах», связана с тем, что данная НС работает на принципе регионов. Регионы имеют небольшую локализацию, в отличие от данных классов. На основании этого можно сделать вывод, что архитектура нейронной сети Mask R-CNN не подходит для распознавания объектов, которые имеют большую локализацию относительно изображения, так как сеть плохо выделяет пространство признаков из объектов, которые имеют большой размер и сложное строение.