

ОБНАРУЖЕНИЕ АВТОМОБИЛЬНЫХ АВАРИЙ В РЕАЛЬНОМ ВРЕМЕНИ С ПРИМЕНЕНИЕМ НЕЙРОСЕТИ АРХИТЕКТУРЫ YOLOV3

*В.Г. Ротарь, к.т.н., доц. ОИТ ИШИТР,
С.И. Беляев, студент гр. 8К92
Томский политехнический университет
E-mail: sib5@tpu.ru*

Введение

Согласно данным ВОЗ, всего в мире происходит 1,35 миллиона смертей и около 20-30 миллионов травм, вызванных автомобильными авариями [1]. Традиционные методы отслеживания автомобильных аварий практически неприменимы на практике по причине низкой точности и эффективности. Внедрение современных технологий компьютерного зрения в сферу регулирования дорожного движения может улучшить ситуацию.

Поэтому целью моей работы являлось написание программного обеспечения, которое дает возможность отслеживать автомобильные аварии в видеопотоке в режиме реального времени, что позволило бы автоматически вызывать экстренные службы на место происшествия для оказания помощи пострадавшим.

Решение задачи можно разбить на три основных этапа: выбор алгоритма компьютерного зрения, создание тренировочного и тестового набора с размеченными данными для соответственно обучения и тестирование алгоритма, и непосредственно проверка работы алгоритма на тестовой выборке.

Ход работы

Современные алгоритмы уже успешно применяются для распознавания автомобилей с помощью классификаторов (например, метод Виолы-Джонса [1]), однако обнаружение ДТП является более сложной задачей ввиду большого разнообразия возможных аварий. Чтобы справиться с этим необходимо применять методы машинного обучения, в частности разнообразные нейронные сети. Некоторые решения задачи предотвращения аварий с применением рекуррентных нейронных сетей достигают хорошей эффективности в решении данной задачи: до 56.14% точности предсказания аварий за 2 секунды до их происшествия [2].

Однако, в настоящее время самыми популярными нейронными сетями для сегментации объектов являются различные вариации сетей YOLO по причине их сверхвысокой скорости обработки видео (You only look once) [3]. YOLO – современная нейронная сеть, позволяющая распознавать объекты на фото и в видеопотоке. Основное её отличие от других активно используемых архитектур, это ее высокая скорость. Архитектуры семейства YOLO в самом деле быстрее R-CNN, что позволяет распознавать объекты в режиме реального времени.

В качестве материала для создания обучающего набора данных были взяты фотографии различных автомобильных аварий из открытых источников. Один из основных аспектов работы, обуславливающих ее сложность, это большое разнообразие автомобильных аварий, вследствие чего обучающая выборка данных должна быть достаточного размера. Все изображения были помечены вручную с помощью утилиты LabelImg [4].

Так как невозможно собрать достаточное количество изображений всех возможных аварий и установить правильные веса для нейросети, было решено использовать веса для классификации автомобилей для последующего применения transfer learning, при этом обучая распознавать аварии лишь самые последние слои нейросети. Это позволило избежать переобучения, вызванного малым количеством данных, размеченных вручную.

Инструментом обучения и непосредственно обнаружения аварии был выбран язык Python с библиотекой OpenCV, предоставляющей все необходимые инструменты установления параметров нейросети. Для обучения классификатора была выбрана платформа облачных вычислений Google Collab, предоставляющая в пользование GPU Tesla K80. Обучение нейросети на основе данных велось порядка пяти часов. В результате был получен классификатор, подходящего для тестирования на тестовой выборке.

Тестирование алгоритма обнаружения аварий

Для оценки работы алгоритма он был протестирован на фото и видео различного качества и степени зашумленности. Само тестирование алгоритма было проведено на CPU. Количество кадров в видео варьировалось от 20 до 30 кадров в секунду что создания максимально правдоподобных условий.

На выходе получалось видео либо изображение с соответствующей рамкой в случае обнаружения автомобильной аварии. На рисунке 1 представлен пример обнаружения аварии на кадре из видео.



Рис.1 Обнаружение аварии на кадре из видео

Заключение

Таким образом, в результате тестирования алгоритма в реальном видеопотоке была достигнута точность в 62% точность определения аварий среди всей тестовой выборки. Алгоритм плохо срабатывает в случае зашумленных кадров аварий, в случае низкого разрешения видео, а также низкой частоты кадров. Предположительно с низким качеством кадров поможет справиться предварительная обработка изображений генеративно-сопоставительными нейросетями, но в таком случае очень значительно пострадает скорость распознавания видео. При обработке отдельных кадров в хорошем качестве алгоритм может достигать порядка 90% распознаваний автомобильных аварий.

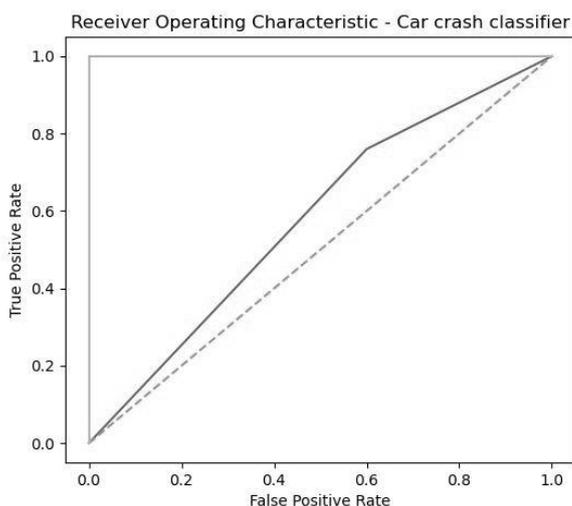


Рисунок 2. ROC-кривая работы алгоритма на тестовых данных

Список использованных источников

1. Сайт Всемирной Организации Здравоохранения. [Электронный ресурс]. – URL: <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries> (дата обращения: 13.02.2021).
2. Chan F.H., Chen Y.T., Xiang Y., Sun M. Anticipating accidents in dashcam videos in Computer Vision. –ACCV–2016. – Springer. – P. 136–153.
3. Redmon J., Divvala S., Girshick R., Farhadi A. You Only Look Once: Unified, Real-Time Object Detection. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – Las Vegas, NV, USA. – 2016. – P. 779-788.
4. LabelImg [Электронный ресурс]. – URL: <https://github.com/tzutalin/labelImg> (дата обращения 13.02.2021).