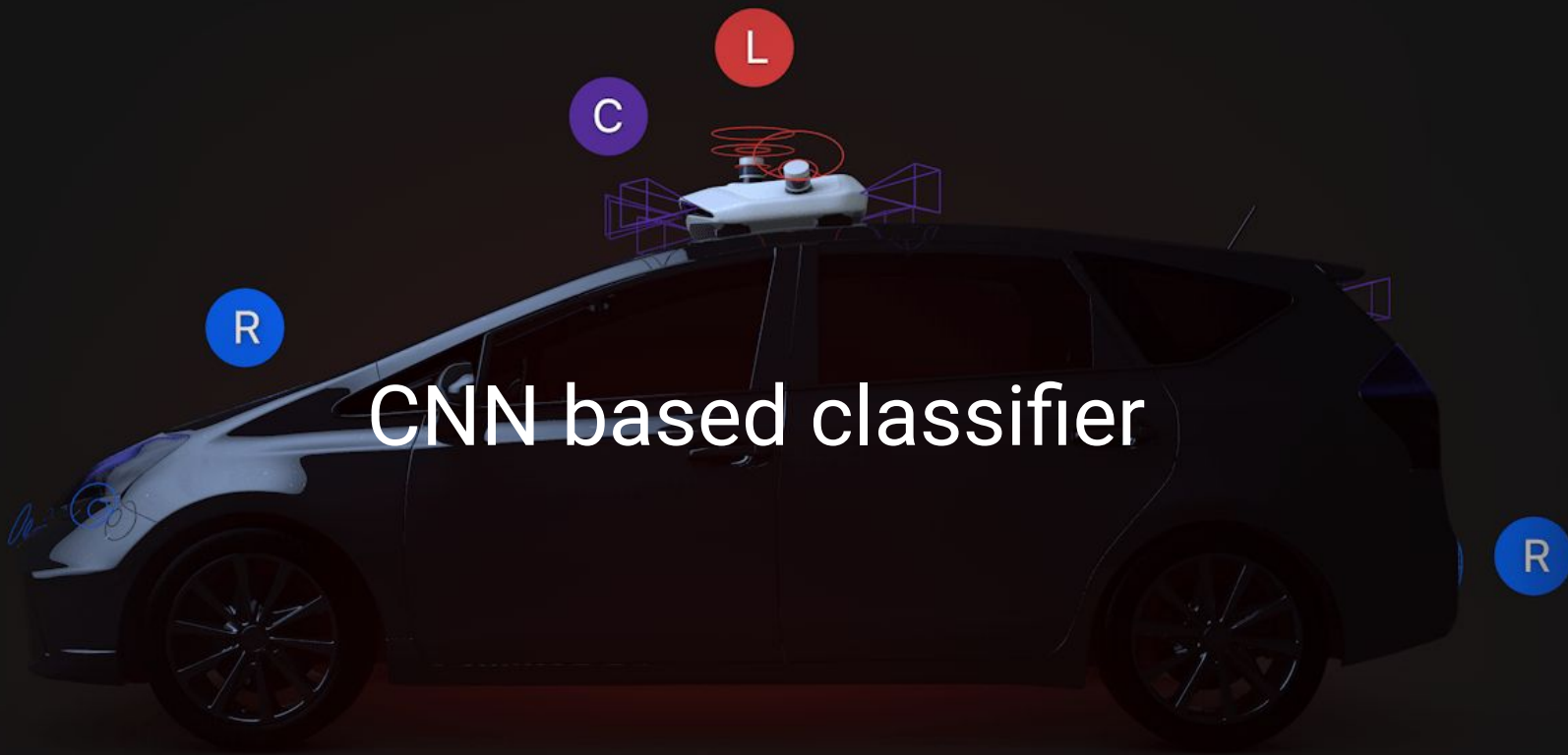


Детектирование объектов в задаче управления беспилотным транспортным средством



CNN based classifier

L – LIDAR

R – RADAR

C – Camera

Архитектура модели

Convolution Layer

Convolution Layer
BatchNorm
ReLU

*Block - repeat
3 times*

Convolution Layer

Dropout 0.2

Fully Connected

Kaggle блокнот

В качестве фреймворка был выбран PyTorch, для вычисления метрики accuracy использовался модуль PyTorch Lightning.

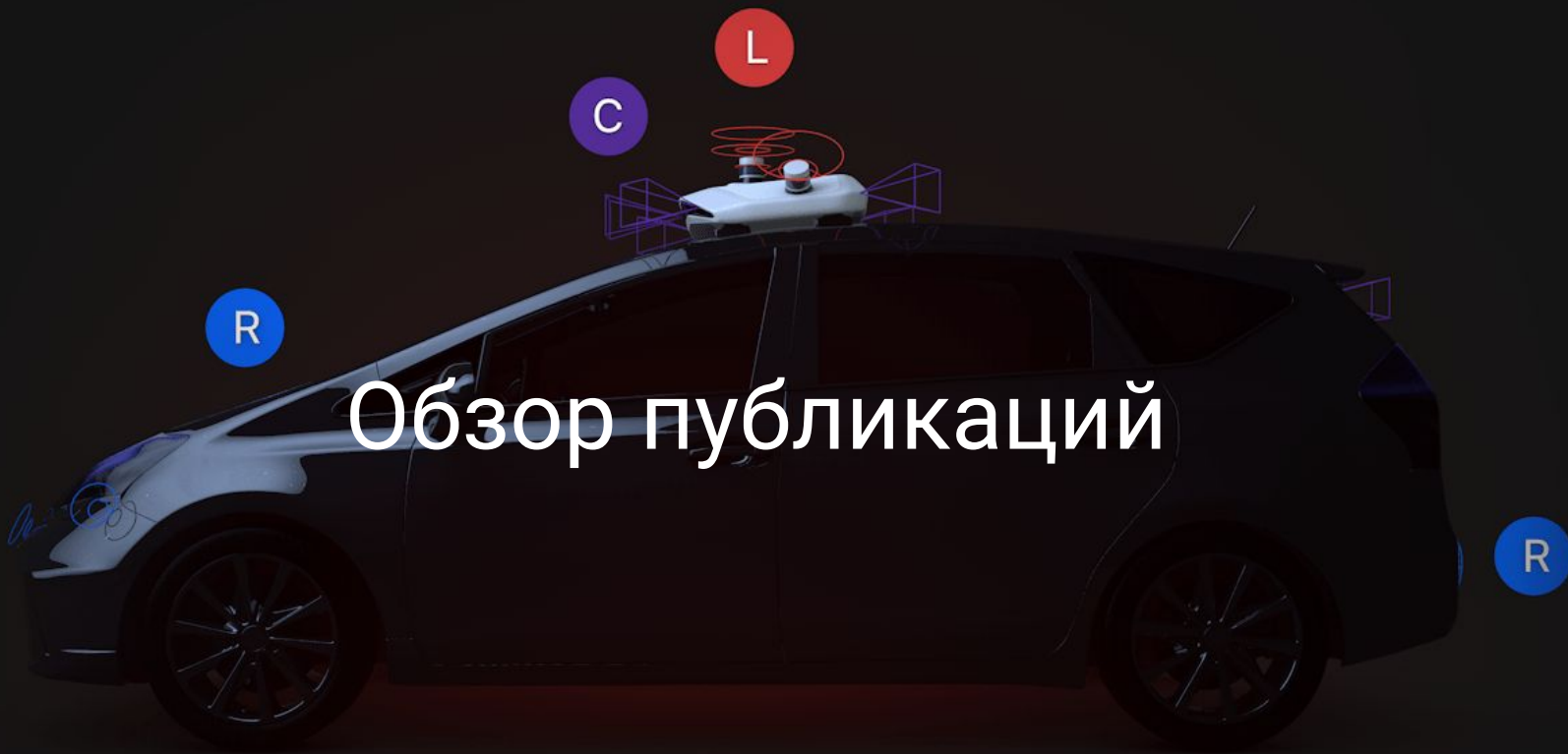
В ходе работы были проведены тестовые запуски с меньшим количеством конволюционных слоев, добавлением Dropout в основной блок и заменой BatchNorm на InstanceNorm.

Предпринимались попытки заменить оптимизационную функцию Адама на SGD, а активационную функции ReLU на Leaky ReLU. Все описанные стратегии были также протестированы на нескольких гиперпараметрах.

Лучший результат на валидационном датасете:
90 эпох, accuracy 0.9902, loss: 0.0261.

На 100 эпохе loss на валидационном датасете начинает расти, что может свидетельствовать о переобучении модели, поэтому оптимальное количество эпох для обучения модели – 90.

Обзор публикаций



L – LIDAR

R – RADAR

C – Camera

Context R-CNN: Long Term Temporal Context for Per-Camera Object Detection

Обычно, частота дискретизации фреймов на камерах статического наблюдения нерегулярна, в связи с чем к моделям предъявляется дополнительное требование устойчивости. В статье авторы предлагают метод, который использует немаркированные кадры с камер статического наблюдения. Подход позволяет представленной в статье модели Context R-CNN индексировать в банк долговременной памяти, созданные для каждой камеры базовые и агрегированные контекстные особенности из других фреймов, что ведет к повышению точности обнаружения объектов в текущем кадре. Авторы применили модель к задачам обнаружения видов животных с помощью камер-ловушек и распознаванию автомобилей на камерах наружного наблюдения. В обоих случаях результаты применения модели превзошли базисное решение.

Сотрудники Калифорнийского Технологического университета и компании Google, 2020

Цитирований: 12

Probabilistic Oriented Object Detection in Automotive Radar

Автомобильные радары используют алгоритмы цифровой обработки сигналов для необработанных данных, редких выводов радара, которые не предоставляют информацию о размере и ориентации объектов. В статье авторы предлагают алгоритм обнаружения радарных объектов, который принимает данные радара в их необработанном тензорном представлении и размещает вероятностно ориентированные ограничивающие прямоугольники вокруг обнаруженных объектов. Авторы статьи создали новый мультимодальный набор данных с 102544 кадрами и разработали конвейер обнаружения транспортных средств – лучший результат достиг 77,28% AP. Предполагается, что это первая попытка обнаружения объектов с помощью обычных угловых автомобильных радаров, используя необработанные данные.

Сотрудники компании XSense.ai, 2020

Цитирований: 2

Distant Vehicle Detection Using Radar and Vision

Возможность автономных транспортных средств оценить скорость сближения с другими транспортными средствами обеспечивает безопасное движение на дорогах. Методы обнаружения объектов, основанные на CNN показывают хорошие результаты на существующих наборах данных, таких как KITTI. Однако, такие модели плохо справляются с распознаванием небольших (удаленных объектов). В статье авторы показывают, что включив данные радара, можно повысить точность распознавания на таких сложных случаях. Авторы также представляют эффективный автоматизированный метод обучения генерации данных с помощью камер разного фокусного расстояния.

Сотрудники Оксфордского института робототехники, 2019

Цитирований: 39

A Deep Learning-based Radar and Camera Sensor Fusion Architecture for Object Detection

Изображения, используемые для распознавания объектов менее эффективны, если они получены ночью или в плохих погодных условиях. Авторы представляют в статье подход который повышает точность распознавания 2D объектов путем объединения данных с камеры и проецируемых разреженных данных радара в слоях нейросети. Представленная CRF-Net автоматически определяет оптимальные пропорции для объединения данных полученных с разных источников. Вдохновленные Dropout, авторы представляют BlackIn – стратегию, которая помогает сфокусировать обучение на данных с сенсора определенного типа. На примере двух датасетов авторы показывают, что результаты использование такой модели превосходят State of the art результаты сетей, основанных на распознавании только одних изображений.

Сотрудники Мюнхенского Технологического университета и компании Automotive Technology, 2019

Цитирований: 20

Deep Multi-modal Object Detection and Semantic Segmentation for Autonomous Driving: Datasets, Methods, and Challenges

Своему развитию область беспилотных транспортных средств в последние годы во многом обязана глубокому обучению. Данные, которые используются для построения моделей приходят с камер, LiDAR и радаров. Многие методы глубокого обучения предполагают необходимость объединения данные из разных источников в процессе обучения, однако отсутствует единое мнение на счет подходящей сетевой архитектуры, дизайна и ответы на вопросы: что, с чем и как объединять, остаются открытыми. В данном обзоре авторы предприняли попытку систематического обобщения методологий и обсуждения проблем: глубокого мультимодального обнаружения объектов и семантической сегментации в области исследования беспилотных автомобилей.

Сотрудники Технологического института Карлсруэ и компании Robert Bosch GmbH, 2020

Цитирований: 114