

Министерство образования Республики Беларусь
Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК 004.032.26

Козлова
Оксана Владимировна

Программное средство глубинного обучения нейронных сетей

АВТОРЕФЕРАТ

на соискание степени магистра информатики и вычислительной техники
по специальности 1-40 81 02 Технологии виртуализации и облачных
вычислений

Научный руководитель

Лукашевич Марина Михайловна
кандидат технических наук, доцент

Минск, 2019

ВВЕДЕНИЕ

Глубинное обучение нейронных сетей – быстро развивающаяся часть машинного обучения, которая позволяет решать задачи, связанные с большими объемами данных в любых отраслях. Системы глубокого обучения нашли применение в таких областях, как компьютерное зрение, распознавание речи, обработка естественного языка, аудиораспознавание, биоинформатика, где для ряда задач были продемонстрированы существенно лучшие результаты, чем ранее.

Глубинное обучение нацелено на реализацию проектов, которые, как правило, требуют многоуровневой структуры обучения. Эта концепция основывается на способности глубокого обучения самостоятельно изучать данные без какого-либо вмешательства человека. Его цель – решать проблемы без каких-либо установленных правил. Наиболее часто глубинное обучение применяется для решения задач, связанных с классификацией, детекцией или сегментацией объектов на изображениях.

Глубинное обучение является одной из самых важных стратегических тенденций. Оно влияет на такие технологии как «Интернет вещей», машинное зрение и другие. Несомненно, что глубокое обучение – это тенденция, которая становится важной технологией, используемой в различных отраслях.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования

Одной из важных тенденций в автомобильной промышленности в настоящее время является разработка автомобилей с самостоятельным управлением (автономные транспортные средства или интеллектуальные транспортные средства).

Степень, в которой транспортному средству предоставляется свобода для автономного вождения, называется уровнем автоматизации. Всего существует 5 уровней. Начиная с уровня ноль (полностью управляется водителем) до пятого уровня, когда автомобиль представляет собой полностью автономную машину для самостоятельного вождения.

Текущее состояние техники включает только уровень 2, в котором управление рулем и ускорением может управляться автомобилем, но водитель должен быть готовым принять управление. Наилучшим представителем данного уровня является автопилот Tesla.

Работу автопилота можно разделить на 4 части: зондирование, восприятие, прогнозирование и планирование.

На этапе зондирования данные собираются от датчиков, стадия восприятия отвечает за использование этой информации для создания понимания окружающей среды с помощью компьютерного зрения. На этой стадии активно используется глубинное обучение нейронных сетей.

Одной из задач, которые решаются на стадии восприятия, является детекция пешеходов. Сложность задачи заключается в том, что необходимо различать людей, которые находятся в движении и возможно собираются пересекать проезжую часть, от людей, которые стоят на тротуаре.

Степень разработанности проблемы

Соревнование алгоритмов машинного обучения в определении пешеходов проводятся в рамках The Multiple Object Tracking Benchmark с 2015 года. В соревновании существуют следующие категории: детекция пешеходов на изображениях, детекция 3D объектов и сегментация.

В 2017 году победителем стал детектор UNV_Det с оценкой Multi-object detection accuracy 81,9.

Также необходимо отметить работы Hailong Li «Pedestrian detection based on deep learning model» и Katleho Masita «Pedestrian Detection Using R-CNN Object Detector».

Данные исследования показывают, что система обнаружения пешеходов,

основанная на глубоком обучении, превосходит традиционные методы детекции.

Цель и задачи исследования

Целью диссертации является разработка программного модуля глубинного обучения, который позволит детектировать и классифицировать объекты.

В соответствии с целью были сформулированы **следующие задачи**:

- провести анализ платформ и библиотек, используемых в глубинном обучении;
- провести анализ архитектур нейронных сетей, подходящих для детектирования и классификации объектов на изображениях;
- разрабатываемый модуль должен классифицировать объекты в соответствии с классами, предоставляемыми «The Multiple Object Tracking Benchmark», главной задачей которого является детекция пешеходов;
- разрабатываемый модуль должен быть обучен на базе изображений MOT17 Challenge;
- результатом работы модуля должно являться изображение, на котором выделены границы объектов и подписан класс найденных объектов.

Объектом исследования работы выступают цифровые изображения.

Предметом исследования являются архитектуры и методы глубинного обучения нейронных сетей.

Область исследования. Содержание диссертационной работы соответствует образовательному стандарту высшего образования второй ступени (магистратуры) специальности 1-40 81 02 «Технологии виртуализации и облачных вычислений».

Теоретическая и методологическая основа исследования

В основу диссертации легли результаты известных исследований российских и зарубежных исследователей в области глубинного обучения нейронных сетей.

Эксперименты с нейронными сетями проводились с использованием библиотеки машинного обучения Tensor Flow на базе облачного сервиса Google Colaboratory.

Информационная база исследования для обучения нейронных сетей сформирована на основе датасета MOTChallenge.

Научная новизна диссертационной работы заключается в разработке программного модуля на основе алгоритма Faster R-CNN, который решает задачу обнаружения пешеходов и сопутствующих классов на изображениях.

Основные положения, выносимые на защиту

1. Анализ подходов детектирования объектов на изображениях, которые в своей основе используют глубинное обучение нейронных сетей.
2. Структура сверточных нейронных сетей, которые подходят для детекции пешеходов.
3. Экспериментальные закономерности, полученные при обучении детектора Faster R-CNN с использованием различных сверточных нейронных сетей. Выявленные зависимости позволяют настроить детектор для качественного определения объектов на изображениях.

Теоретическая значимость диссертации заключается в том, что проведен анализ сверточных нейронных сетей и подходов детекции объектов.

Практическая значимость диссертации состоит в том, что разработан программный модуль глубинного обучения нейронных сетей, который позволяет определять пешеходов и сопутствующие объекты (транспортные средства, окклюдеры) на изображениях.

Апробация и внедрение результатов исследования

Результаты исследования были представлены на 54-ой научной конференции аспирантов, магистрантов и студентов БГУИР и на пятой международной научно-практической конференции «BIG DATA and Advanced Analytics».

Публикации

Основные положения работы и результаты диссертации изложены в двух опубликованных работах.

Структура и объем работы

Структура диссертационной работы обусловлена целью, задачами и логикой исследования. Работа состоит из введения, четырех глав и заключения, библиографического списка и приложений. Общий объем диссертации – 101 страница. Работа содержит 14 таблиц, 64 рисунка. Библиографический список включает 30 наименований.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении рассмотрено современное состояние проблемы детекции пешеходов на изображении, определены основные направления исследований, а также дается обоснование актуальности темы диссертационной работы.

В общей характеристике работы сформулированы ее цель и задачи, показана связь с научными программами и проектами, даны сведения об объекте исследования и обоснован его выбор, представлены положения, выносимые на защиту, приведены сведения о личном вкладе соискателя, апробации результатов диссертации и их опубликованность, а также, структура и объем диссертации.

В первой главе рассматриваются существующие подходы детекции объектов на основе глубинного обучения и сверточных нейронных сетей. Также приведен обзор фреймворков и библиотек для машинного обучения.

Во второй главе описана работа алгоритма детекции Faster R-CNN.

В третьей главе представлены архитектуры сверточных нейронных сетей, которые используются совместно с Faster R-CNN.

В четвертой главе описаны эксперименты, проводимые с детектором Faster R-CNN, в которых изменялась сверточная нейронная сеть детектора.

В приложении приведен исходный код, разработанного программного модуля и презентация.

ЗАКЛЮЧЕНИЕ

В рамках данной работы был проведен анализ детекторов и сверточных нейронных сетей для решения задачи детекции пешеходов на изображениях. Для решения данной задачи был выбран алгоритм детекции Faster R-CNN.

После анализа сверточных нейронных сетей было решено проводить эксперименты для сетей ResNet50, ResNet101 и Inception V2. Целью экспериментов являлась определение наиболее подходящей сети для детектора Faster R-CNN, который обучался на датасете MOTChallenge. В результате экспериментов было установлено, что для решения задачи детекции пешеходов с помощью детектора Faster R-CNN наиболее подходящей сверточной нейронной сетью является Inception V2.

В процессе обучения сверточных нейронных сетей было выявлено, что в датасете аннотированы не все объекты определяемых классов. Данная особенность искажает значения метрик, так как при оценке качества объекты, который детектор правильно определил, интерпретируются как ошибка детекции.

Также был произведен анализ фреймворков и библиотек для глубинного обучения. Было решено использовать Tensor Flow, так как данная библиотека позволяет решать различные задачи машинного обучения, имеет инструменты для оценки качества и большое комьюнити.

На основе экспериментальных данных и анализа предметной области был разработан программный модуль, который позволяет детектировать пешеходов на изображениях. В программном модуле используется детектор Faster R-CNN со сверточной нейронной сетью Inception V2.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1. Козлова, О. В. Анализ основных архитектур нейронных сетей глубокого обучения / О. В. Козлова // Компьютерные системы и сети: материалы 54-й научной конференции аспирантов, магистрантов и студентов, Минск, 23 – 27 апреля 2018 г. / Белорусский государственный университет информатики и радиоэлектроники. – Минск, 2018. – С. 45.

2. Козлова О.В., Куница Е.Ю., Лукашевич М.М. U-net для решения задачи сегментации медицинских изображений / BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сб. материалов V Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 13–14 марта 2019 года). В 2 ч. Ч. 2 / редкол. : В. А. Богуш [и др.]. – Минск : БГУИР, 2019. – 379 с. – С. 295.