

## ПРИМЕНЕНИЕ БОЛЬШИХ ДАННЫХ В АВТОМАТИЗАЦИИ ДОРОЖНОГО ДВИЖЕНИЯ.

*Пихтовникова М.Т.<sup>1</sup>, студент гр.153502, Захаренко Н.В.<sup>2</sup>, студент гр.153501*

*Белорусский государственный университет информатики и радиоэлектроники г. Минск,  
Республика Беларусь*

*Марков А.Н. – старший преподаватель каф. информатики*

**Аннотация.** В данной статье рассматривается сущность и значимость использования *big data* в сфере автоматизации дорожного движения. Обсуждаются основные принципы, технологии и преимущества применения аналитики больших данных для улучшения эффективности управления дорожным движением, повышения безопасности на дорогах и оптимизации транспортной инфраструктуры. Продемонстрировано, как аналитика данных, алгоритмы машинного обучения и Интернет вещей могут использоваться для прогнозирования транспортного потока, создания точных данных о дорожных заторах в режиме реального времени и перенаправления автомобилей с помощью навигации на менее загруженный маршрут.

**Ключевые слова:** аналитика данных, автоматизация дорожного движения, машинное обучение, GPS, анализ изображений, интеллектуальная система управления трафиком, IoT.

**Введение.** С постоянным ростом числа автотранспортных средств на дорогах возникают серьезные проблемы с управлением дорожным движением. Такие проблемы как *заторы, аварии, несоблюдение правил дорожного движения и другие факторы приводят к ухудшению безопасности и эффективности транспортного потока.*

Традиционные методы контроля и регулирования становятся недостаточно эффективными для обеспечения безопасности и оптимизации транспортного потока.

Для *предсказания и предотвращения аварий, улучшения координации светофоров, оптимизации маршрутов, управления потоками транспорта, прогнозирования транспортного потока, выявления аномалий на дороге и оптимизации работы инфраструктуры* используют новые разработки в области машинного обучения и алгоритмы анализа данных, собранных на дорогах с различных источников таких как *видеокамеры, системы распознавания, радары, GPS-трекеры и другие.*

Благодаря быстрому развитию коммуникационных и сенсорных технологий, недорогим и эффективным датчикам, повышению эффективности хранения и извлечения данных, а также недорогому хранению больших объемов данных, извлекать и использовать данные стало проще. Однако, основной проблемой в анализе данных является сбор релевантных и полезных данных для разработки решения. В модель данных необходимо загружать постоянно обновляемые данные, а методы прогнозирования, генерируемые алгоритмами, должны обеспечивать предоставление корректных отчетов на основе этих постоянно растущих данных.

*Цель данной статьи — рассмотреть, как использование методов анализа больших данных позволяет не только автоматизировать сбор информации о дорожном движении, но и обучать модели, способные предсказывать и оптимизировать его поведение. Будут рассмотрены различные аспекты применения Big Data в автоматизации дорожного движения, начиная с сбора данных из различных источников, проходя через их хранение и обработку, и заканчивая применением аналитики для оптимизации транспортного потока и повышения безопасности.*

**Автоматизированные методы сбора и хранения данных.** Одним из ключевых аспектов применения больших данных в автоматизации дорожного движения является обучение моделей на основе собранных данных.

Данные GPS, собираемые с автомобилей – весьма полезные при разработке модели данных. Датчик GPS (глобальная система позиционирования) определяет точное местоположение автомобиля. Зная местоположение всех автомобилей, можно предвидеть, возникнут ли заторы на дорогах. Эти данные особенно ценны для определения скорости или плотности движения в определенном месте. Плотность дорожного движения оценивается, сравнением местоположения данного автомобиля с количеством автомобилей, присутствующих в радиусе 100 метров от этого места. Скорость транспортного средства также оказывает существенное влияние на это. Еще одним источником полезных данных являются камеры видеонаблюдения, установленные на маршруте.

*Эффективный сбор данных позволяет получить информацию о текущем состоянии дорожного потока, аномалиях, аварийных ситуациях и других важных параметрах, необходимых для принятия решений и оптимизации инфраструктуры. Данные, полученные из видеокамер, радаров,*

*трекеров, устройств пользователей и тому подобных, должны перед тем, как поступить на вход в обучаемую модель, обработаться.*

*Например, системы видеонаблюдения и датчики на дорожных знаках и светофорах могут непрерывно собирать информацию о транспортных потоках и дорожных условиях при этом обновляя свои DataSet и увеличивая массив данных все больше и больше для обучения модели. Чем больше объем данных, поступающих в обучаемую модель, тем точнее полученный результат прогноза.*

*Для эффективного обнаружения объектов в реальном времени используются одноступенчатые детекторы и двухступенчатые детекторы. Одноступенчатые детекторы часто работают быстро и могут предсказать границы объектов, а также классов за один запуск сети. YOLO и SSD - два хорошо известных одноступенчатых детектора. Эти проекты особенно хорошо работают, когда целевые объекты занимают значительную часть изображения. [1]*

*Примером такого рода данных является знаменитый набор данных обнаружения транспортных средств UA-DETRAC. Еще одним примером хорошего компромисса между скоростью и точностью является архитектура YOLO v2, которая была оптимизирована для распознавания транспортных средств с помощью кластеризации привязки, нормализации дополнительных потерь и метода многоуровневого объединения объектов. [2]*

*Датчики, установленные в Интеллектуальные Транспортные Системы (Intelligent Transportation Systems), собирают такие данные, как скорость движения, плотность транспортных средств, транспортные потоки и время в пути. Дорожные датчики (например, инфракрасные и микроволновые детекторы) эволюционировали для сбора, расчета и передачи данных о дорожном движении. [3]*

*Сбор данных с датчиков делится на три категории: данные о дорогах, данные о плавающих автомобилях и данные о обширной территории.*

*Данные о дорогах включают в себя данные, связанные с состоянием и параметрами дорожного покрытия, такими как температура, влажность, прочность и деформация дорожных покрытий. Датчики для сбора таких данных могут включать в себя термометры, гидрометры, датчики давления и т.д. Благодаря последним достижениям в области технологий постепенно становятся доступными придорожные датчики следующего поколения, в том числе ультразвуковые и акустические сенсорные системы, магнитные детекторы транспортных средств, инфракрасные системы, системы обнаружения света и дальности (LIDAR), а также системы обработки и обнаружения видеоизображений. Данные о плавающих автомобилях в первую очередь относятся к данным, связанные с транспортными средствами и их движением, такие как скорость, ускорение, торможение, уровень топлива и т. д. Они собираются с помощью различных датчиков, включая GPS, датчики скорости, ускорения, уровня топлива и т. д. Определенные встроенные датчики предоставляют надежные и эффективные данные для выбора и оценки маршрута. Популярные технологии датчиков FCD включают автоматическую идентификацию транспортных средств (AVI), распознавание номерных знаков (LPR) и транспондеры, такие как зондирование автомобилей и электронные дорожные знаки. Данные по обширной территории относятся к данным о транспортном потоке, полученным на большой территории с использованием различных методов мониторинга датчиков, включая фотограмметрическую обработку, запись звука, обработку видео и космический радар.*

*Широко используются данные, собранные из открытых источников служб такси. Эти данные, касающиеся автомобилей и маршрутов движения, используемых в приложениях, являются ценным ресурсом для создания моделей данных и прогнозирования трафика. Использование таких данных улучшает качество прогнозирования трафика, так как водители в основном следуют по маршруту, указанному в приложении, что обеспечивает регулярное обновление данных. Это позволяет получать информацию об изменениях в городе и обновлениях трафика в режиме реального времени. Такие данные активно используются для обучения моделей, поскольку некоторые паттерны в движении транспорта могут повторяться ежедневно, особенно для пассажиров, которые регулярно пользуются такси для поездок на работу или в другие места.*

*После сбора данных, они передаются для дальнейшей обработки, и для этого существует несколько технологий, которые позволяют эффективно и быстро передавать данные с распознавательных устройств на дорогах. Вот 2 самые ключевые из них: передача данных по беспроводной сети, передача с использованием технологии IoT.*

*Использование беспроводных сетей позволяет передавать данные с датчиков и камер на центральные серверы для дальнейшей обработки и анализа. Это выполняется с использованием технологий, таких как Wi-Fi, Bluetooth, сотовая связь или специализированные протоколы передачи данных таких как UDP или TCP/IP, которые являются основными протоколами передачи данных в сети Интернет.*

*Технология IoT позволяет подключать различные устройства и сенсоры к сети интернет. Основная идея IoT заключается в том, что все эти объекты могут взаимодействовать между собой и с*

внешними системами, а также принимать самостоятельные решения на основе полученной информации. И так же, как и концепции беспроводной передачи данных, *IoT* для передачи использует различные протоколы передачи данных.

Распределив процесс сбора данных на несколько узлов или устройств, значительно снижается нагрузка на отдельные компоненты сети и обеспечивается более быстрая и эффективная передача данных.

Для хранения данных и моделей, а также для выполнения вычислений в облаке, используются различные облачные сервисы, такие как *Amazon Web Services (AWS)*, *Google Cloud Platform (GCP)* или *Microsoft Azure*. [4]

**Один из эффективных инструментов для этой цели - *Google BigQuery*, облачный сервис аналитики данных, предназначенный для хранения и анализа больших объемов структурированных данных с использованием SQL-запросов.** После сбора данных и их загрузки в *Google BigQuery* проводится предварительная обработка и очистка данных, а также создаются признаки, которые будут использоваться для обучения моделей машинного обучения.

Интеграция с облачными *API* позволяет использовать мощные вычисления и хранилища данных, предоставляемые облачными провайдерами, что может быть полезно для масштабирования системы, а также обработки больших объемов данных.

Эффективным инструментом для обработки потоков больших данных, полученных с дороги, является *Kafka*. Благодаря встроенным механизмам *Kafka* таким как *Apache Storm*, *Apache Flink*, *Spark Streaming*, которые позволяют выполнять различные операции с данными, такие как фильтрация, агрегация, преобразование и анализ, в реальном времени, что особенно важно для мониторинга и управления дорожным движением, оценка записей является масштабируемой, т. е. может масштабироваться до большого количества источников записей, одновременно отправляющих записи с высокой скоростью, и надежной, т. е. может допускать аппаратные сбои без потери записей.

**Обучение моделей на базе Big Data.** Одним из ключевых аспектов применения больших данных в автоматизации дорожного движения является обучение моделей на основе собранных с дороги данных. Эти данные используются для обучения алгоритмов машинного обучения. *Обучение моделей на основе Big Data с использованием современных инструментов машинного обучения, таких как Keras и TensorFlow, является ключевым аспектом применения больших данных в автоматизации дорожного движения.* [5]

*Keras* — это высокоуровневый интерфейс для построения и обучения нейронных сетей, который работает поверх более низкоуровневых библиотек, таких как *TensorFlow*. *TensorFlow*, в свою очередь, является мощным фреймворком для создания и обучения глубоких нейронных сетей. С их помощью создано много моделей для автоматизации дорожного движения, прогнозирования аварий, анализа и оптимизации маршрутов и тп. Ниже представлены 2 самые важные из них.

С использованием *TensorFlow* и *Keras* можно обучать сверточные нейронные сети (*CNN*), которые широко используются для обнаружения и классификации объектов на дороге, таких как автомобили, пешеходы, велосипедисты и транспортные средства общественного транспорта, полученных с камер видеонаблюдения. Она разбивает изображение на сетку (*grid*) и для каждой ячейки сетки предсказывает ограничивающие рамки (*bounding boxes*) и вероятности классов. Пример как это работает изображен на рисунке 1. Эта нейронная сеть используется для автоматического анализа дорожного движения, выявления нарушений правил и предотвращения аварий. [6]

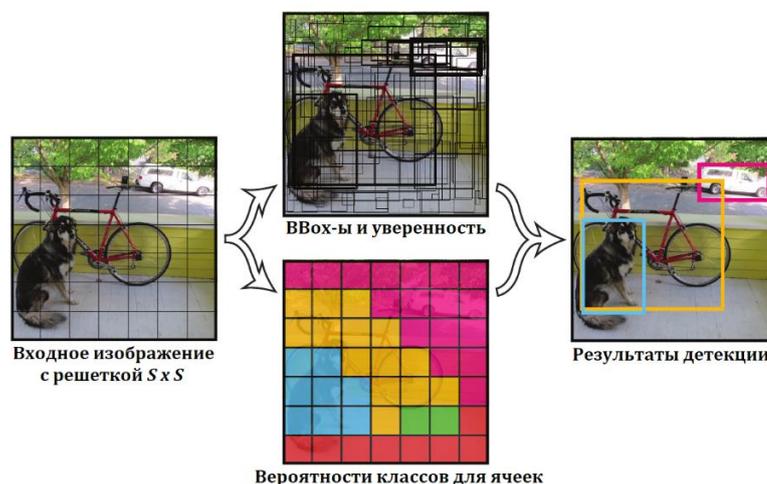


Рисунок 1 – Пример работы сверточной нейронной сети

Рекуррентные нейронные сети, реализованные на базе TensorFlow и Keras с долгой краткосрочной памятью (LSTM) широко применяются для прогнозирования будущего трафика на основе исторических данных о трафике, погоде, событиях на дороге и других факторах.

На данный момент большинство детекторов объектов основаны на сверточных нейронных сетях (CNN). Например: YOLO и SSD.

Для обучения всех этих моделей необходимо собрать большой массив данных и обработать его, для этого существует множество подходов, о которых пойдет разговор далее.

**Применение аналитики данных для управления дорожным движением в больших городах.** Основываясь на анализе больших объемов данных, предсказываются и предотвращаются аварии, совершенствуется координация светофоров, оптимизируются маршруты, выявляются опасные участки дороги. Это помогает снизить количество ДТП и улучшить общую безопасность на дорогах.

Анализ данных позволяет выявлять паттерны и факторы, способствующие возникновению аварийных ситуаций. На основе этих данных строятся прогностические модели, которые предсказывают вероятность аварий на конкретных участках дороги или в определенные моменты времени.

Что касается улучшения координации светофоров, то с использованием моделей обученных на больших объемах данных можно анализировать трафик и потоки транспорта в реальном времени. В Лос-Анджелесе, городе с самыми сильными пробками в США, уже есть система, использующая большие данные для управления светофорами. Описание из газеты New York Times о том, как это работает: «Магнитные датчики на дорогах на каждом перекрестке отправляют обновленную информацию о транспортном потоке в режиме реального времени по волоконно-оптическим кабелям в бункер под центром Лос-Анджелеса, где Эдвард Ю управляет сетью. Компьютерная система, которая запускает программное обеспечение, разработанное самим городом, анализирует данные и автоматически вносит посекундные корректировки, адаптируясь к меняющимся условиям и используя массив прошлых данных, чтобы предсказать, где может возникнуть перебои в движении, и все это без участия человека». [7]

Так же в городе Ресифи была автоматизирована система работы светофоров. Светофоры оснащены камерами, чтобы повысить эффективность времени, в течение которого они остаются зелеными или красными, в соответствии с движением на улицах, без необходимости использования заранее запрограммированной схемы. С помощью этих умных светофоров управление транспортным потоком осуществляется в режиме реального времени. На рисунке 2 показано, как работает адаптивная система управления городским движением. [8]

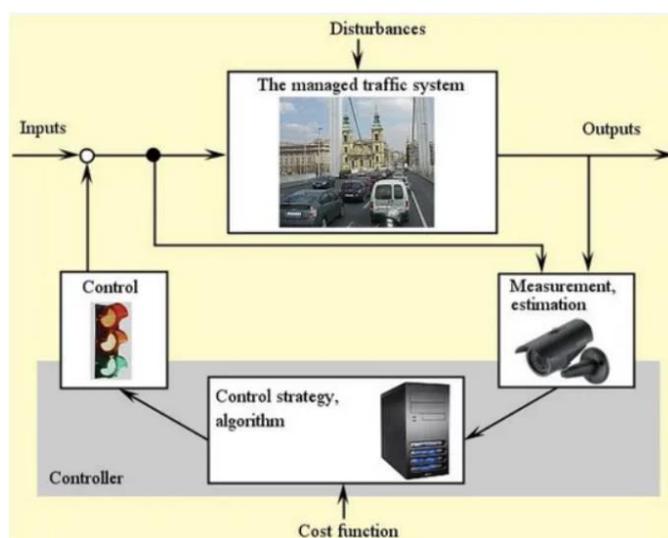


Рисунок 2 – Работа адаптивной системы управления городским движением.

Города по всему миру используют системы для предотвращения аварийных ситуаций с применением технологий машинного обучения и обработки большого объема данных. Примеры проектов, которые существуют для автоматизации дорожного движения и предотвращения аварий: *Array of Things* в Чикаго, *Vision Zero* по всему миру, *Hudson Yards* в Нью-Йорке, в Рио-де-Жанейро, «предотвращение дорожно-транспортных происшествий со смертельным исходом» в Барселоне. [9]

**В рамках проекта *Array of Things*** совместно с *city of Chicago* и альянсом *Loop Alliance* установлены датчики вдоль главных улиц Чикаго. Эти датчики передают данные для более глубокого понимания того, как пешеходы и транспортные средства передвигаются по городу. Градостроители и ученые используют эти данные для повышения безопасности дорожного движения.

Проект ***Vision Zero*** создан для предотвращения аварий со смертельным исходом. Реализация этой программы в различных городах отличается: в некоторых городах основное внимание уделяется датчикам сигналов светофора, в других - проведению пространственного анализа мест, где происходят типы аварий, и определению мест для размещения выделенных велосипедных дорожек.

Что касается проекта ***Centro de Operacoes Prefeitura Do Rio***, то в результате совместного предприятия города местное самоуправление бразильского города располагает автоматизированной системой, содержащей данные более чем от 30 различных агентств со всего города. Используя эти данные, они не только создают отчеты о том, что в настоящее время происходит в городе, но и строят прогностические модели. Эта универсальная модель имеет множество применений по всему городу, включая множество ценной информации о дорожном движении.

Для достижения цели Европейского союза по снижению смертности в результате дорожно-транспортных происшествий на 50% к 2020 году город Барселона внедрил интеллектуальное решение «предотвращение дорожно-транспортных происшествий со смертельным исходом». Его суть в том, что при прогнозировании учитываются условия окружающей среды на данный момент, а также учитываются данные об авариях в городе, включая количество аварий, где они произошли и даже какая была погода. [10]

*Таким образом, применение больших данных в автоматизации дорожного движения не только способствует повышению безопасности, но и значительно улучшает эффективность транспортных систем, что в конечном итоге приносит пользу как водителям, так и городским органам управления дорожным движением. с помощью анализа больших данных.*

**Заключение.** Применение методов анализа больших данных открывает новые перспективы для автоматизации управления дорожным движением. Обучение моделей на базе *Big Data* позволяет создавать более точные и эффективные алгоритмы прогнозирования и оптимизации транспортного потока, а автоматизация процессов сбора информации значительно снижает затраты на управление дорожным движением и повышает безопасность на дорогах. *Модели машинного обучения, обученные на больших данных, обладают высокой точностью предсказаний и могут выявлять тенденции, незаметные для человеческого анализа, и уже активно применяются и улучшаются для решения различного рода проблем, возникающих на дорогах.*

Одним из главных преимуществ использования больших данных в автоматизации дорожного движения является возможность принятия решений в реальном времени на основе актуальной информации о транспортном потоке. Это позволяет оптимизировать управление светофорами, перенастраивать маршруты общественного транспорта, а также оперативно реагировать на аварийные ситуации.

Дальнейшее развитие и применение этих технологий поможет создать более умные и безопасные городские инфраструктуры, способствуя повышению качества жизни горожан и снижению загруженности дорожной сети.

#### **Список использованных источников:**

1. YOLO, SSD и др. [Электронный ресурс]. – Режим доступа: [http://vbystricky.ru/2020/08/yolo\\_ssd\\_etc.html](http://vbystricky.ru/2020/08/yolo_ssd_etc.html) – Дата доступа: – 12.02.2024.
2. The UA-DETRAC Benchmark Suite. [Электронный ресурс]. – Режим доступа: <https://detrac-db.rit.albany.edu/> – Дата доступа: – 12.02.2024.
3. J. Lopes, J. Bento, E. Huang, C. Antoniou, and M. Ben-Akiva, "Traffic and mobility data collection for real-time applications," in Proc. IEEE Intell. Transp. Syst. (ITSC), Sep. 2010, pp. 216–223
4. Обзор основных функций Google BigQuery и примеры запросов для маркетинг-анализа. [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/articles/511086/> – Дата доступа: 14.02.2024.
5. Документация TensorFlow. [Электронный ресурс]. – Режим доступа: [https://www.tensorflow.org/api\\_docs](https://www.tensorflow.org/api_docs) – Дата доступа: 14.02.2024.
6. Что такое сверточная нейронная сеть? [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/articles/309508/> – Дата доступа: 14.02.2024.
7. To Fight Gridlock, Los Angeles Synchronizes Every Red Light. [Электронный ресурс]. – Режим доступа: [https://www.nytimes.com/2013/04/02/us/to-fight-gridlock-los-angeles-synchronizes-every-red-light.html?\\_r=0](https://www.nytimes.com/2013/04/02/us/to-fight-gridlock-los-angeles-synchronizes-every-red-light.html?_r=0) – Дата доступа: – 15.02.2024.
8. Towards Smart Traffic lights Using big data to improve Urban Traffic. [Электронный ресурс]. – Режим доступа: <https://medium.com/@janduijorge/towards-smart-traffic-lights-using-big-data-to-improve-urban-traffic-ffd98873aa3d> – Дата доступа: – 12.02.2024.
9. 6 Real Examples of Cities Making Roads Smarter – and Safer. [Электронный ресурс]. – Режим доступа: <https://blog.bismart.com/en/traffic-fatalities-prevention-making-roads-smarter> – Дата доступа: – 15.02.2024.
10. Traffic Fatalities Prevention. [Электронный ресурс]. – Режим доступа: <https://bismart.com/en/business-intelligence-solutions/traffic-fatalities-prevention/> – Дата доступа: – 15.02.2024.

## **THE USE OF BIG DATA IN TRAFFIC AUTOMATION**

*Pikhtovnikava M.T.<sup>1</sup>, Zakharanka N.V.<sup>1</sup>,*

*Belarusian State University of Informatics and Radioelectronics<sup>1</sup>, Minsk, Republic of Belarus*

*Markov A.N. – senior lecture at the department of informatic*

**Annotation.** This article examines the essence and importance of using big data in the field of traffic automation. The main principles, technologies and advantages of using big data analytics to improve traffic management efficiency, improve road safety and optimize transport infrastructure are discussed. It demonstrates how data analytics, machine learning algorithms and the Internet of Things can be used to predict traffic flow, *receive accurate traffic data in real time* and redirect cars using navigation to a less busy route.

**Keywords:** data analytics, machine learning, GPS, image analysis, intelligent traffic management system, Internet of things.