

---

## ОБЗОР МЕТОДОВ КОМПЬЮТЕРНОГО ЗРЕНИЯ И АРХИТЕКТУР СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ОЦЕНКИ ПЛОТНОСТИ ОБЪЕКТОВ В ГОРОДСКОЙ СРЕДЕ

**Барская Ирина Николаевна,**  
магистрант МГТУ им. Н.Э. Баумана  
kuznetsovamos@gmail.com

**Иманбеков Кубан,**  
магистрант МГТУ им. Н.Э. Баумана  
kuban.imanbekov@gmail.com

**Кушнир Григорий Леонидович,**  
магистрант МГТУ им. Н.Э. Баумана  
Россия г. Москва  
wolf20boss@gmail.com

### Аннотация

---

Данное исследование рассматривает основные аспекты в области дорожных заторов и плотности транспортного потока. Ключевые элементы представлены в рамках структур глубокого обучения, таких как компьютерное зрение и Vision Transformers. Также подробно анализируются модели оценки плотности транспортных средств, включая их характеристики и классификацию на модели с учителем и без учителя. Данное исследование послужит основой для исследователей, работающих в данной области, и станет отправной точкой для дальнейших исследований.

---

**Ключевые слова:** Глубокое обучение, Компьютерное зрение, Плотность транспортного потока, Заторы на дорогах, Обучение с учителем и без учителя.

---

## AN OVERVIEW OF COMPUTER VISION METHODS AND ARCHITECTURES OF CONVOLUTIONAL NEURAL NETWORKS FOR ESTIMATING THE DENSITY OF OBJECTS IN AN URBAN ENVIRONMENT

**Irina N. Barskaya,**  
master's student of Bauman Moscow State Technical University  
Russia, Moscow  
kuznetsovamos@gmail.com

**Kuban Imanbekov,**  
master's student of Bauman Moscow State Technical University  
Russia, Moscow  
kuban.imanbekov@gmail.com

**Grigory L. Kushnir,**

master's student of Bauman Moscow State Technical University

wolf20boss@gmail.com

Bauman Moscow State Technical University

Russia, Moscow

---

## ABSTRACT

---

This article presents an approach for applying SOLID principles in the development of automated tests. SOLID is a well-known set of design principles (Single Responsibility, Open-Closed, Liskov Substitution, Interface Segregation, Dependency Inversion) that improve the modularity, extensibility, and maintainability of object-oriented systems. The paper discusses how each of these principles can be utilized in test automation to create flexible and robust test code. We provide examples in Python, JavaScript, and C# to illustrate correct applications of each principle in test scenarios. Adhering to SOLID in test architecture leads to cleaner test suites that are easier to extend with new tests and less prone to breakage when production code changes. The advantages of a SOLID-compliant test design – such as improved readability, reusability of test components, and more stable test suites – are highlighted. The article is intended for experienced software testers and developers interested in advanced test automation practices. This research illustrates the main keys if the field of traffic jams and vehicle densities in the streets. the main keys are presented as a part of deep learning structures such as computer visions or vision transformers. The details of models for vehicles densities are also discussed with details and types of models supervised and unsupervised learning. This study will be a key for researchers in this field to start working in this field.

---

**Keywords:** Deep learning, Computer vision, Traffic density, Traffic congestion, Supervised learning.

---

Умные города в настоящее время в значительной степени зависят от технологий искусственного интеллекта и современных интеллектуальных систем, способствующих улучшению предоставления услуг. Одной из наиболее актуальных проблем в умных городах является скопление людей и перегруженность улиц, что может приводить к возникновению пробок. В умных городах не должно возникать необходимости в присутствии сотрудников полиции на улицах для регулирования движения. В данном исследовании проводится обзор литературы, посвященный применению методов компьютерного зрения и глубокого обучения для предсказания плотности транспортных средств на улицах с целью адаптивного управления дорожным движением [1].

В ряде исследований разработали адаптивную модель, способную в реальном времени оценивать количество транспортных средств на городских дорогах с использованием технологий компьютерного зрения [6]. В данной работе предложен автоматический алгоритм обновления фона для детекции транспортных средств в реальном времени, а также адаптированная модель подсчета транспортных средств, основанная на методах виртуального кольца и линии обнаружения. Кроме того, был представлен новый надежный метод детекции для мониторинга дорожной обстановки в реальном времени. Для проведения тестирования был разработан прототип системы, который был установлен на городской дороге.

Для эффективного совершенствования и повышения мер безопасности дорожного движения в Саудовской Аравии появилось значительное количество технологий мониторинга дорожного движения на основе искусственного интеллекта (ИИ) [7]. Однако своевременное обнаружение дорожно-транспортного происшествия (ДТП) играет ключевую роль в повышении скорости реагирования служб управления инцидентами, что, в свою очередь, снижает количество травм, вызванных аварией. В целях повышения уровня безопасности дорожного движения авторы представляют систему обнаружения и оповещения о ДТП в реальном времени, основанную на методах компьютерного зрения. Предложенная архитектура включает три модели, интегрированные в прототипный интерфейс для визуализации работы системы. Первая модель, предназначенная для обнаружения и отслеживания транспортных средств, использует детектор объектов YOLOv5 в сочетании с трекером DeepSORT для отслеживания движения автомобилей путем присвоения каждому из них уникального идентификационного номера (ID). Данная модель достигла средней точности (mAP) 99,2%. Вторая модель классификации дорожно-транспортных происшествий по степени тяжести продемонстрировала mAP 83,3%, применяя алгоритм YOLOv5 для точного выявления и классификации серьезности аварии. В случае тяжелого ДТП система автоматически отправляет экстренное сообщение в ближайший медицинский центр. Третья модель, основанная на алгоритме ResNet152, предназначена для обнаружения возгорания, возникающего вследствие аварии. Данная модель достигла точности 98,9%, при этом система автоматически отправляет сигнал тревоги в пожарную службу при выявлении возгорания. Для повышения производительности и сокращения времени вывода данных авторы использовали инновационную технологию параллельных вычислений, что позволило системе работать в многопоточном и параллельном режиме, снижая общую вычислительную сложность.

Приложения компьютерного зрения в интеллектуальных транспортных системах (ИТС) и автономном вождении (АВ) в последние годы все чаще основываются на архитектурах глубоких нейронных сетей. Несмотря на улучшение производительности на эталонных наборах данных, многие реальные проблемы остаются недостаточно изученными в научных исследованиях. Ряд исследователей провел обширный обзор литературы, посвященный применению компьютерного зрения в ИТС и АВ, а также рассмотрел вызовы, связанные с данными, моделями и сложными городскими условиями. Проблемы, связанные с данными, включают сбор и разметку обучающих выборок, их соответствие реальным условиям, наличие системных искажений в данных, необходимость обработки больших объемов информации, а также вопросы конфиденциальности [7].

Модели глубокого обучения зачастую оказываются слишком сложными для выполнения в реальном времени на встроенных устройствах, обладают низкой интерпретируемостью и обобщающей способностью, а также трудно тестируются в реальных условиях. Сложные городские транспортные среды характеризуются нерегулярным освещением, частичными перекрытиями объектов, а также различными условиями установки камер наблюдения (различные углы, загрязнение объективов, колебания от ветра). Кроме того, транспортный поток в городах является высокоразнородным, что сопровождается нарушениями правил дорожного движения и сложными взаимодействиями между транспортными средствами в условиях высокой плотности движения [4].

Некоторые из приложений, подверженные этим проблемам, включают оценку транспортного потока, обнаружение заторов, восприятие окружающей среды в автономном вождении, анализ взаимодействий транспортных средств и применение граничных вычислений для практической реализации. В исследовании также рассматриваются

возможные способы решения указанных проблем с акцентом на практическое внедрение технологий.

Структура программной модели будет состоять из следующих основных компонентов:

Входные данные – это изображения или видеозаписи, полученные с камер, установленных в определенных местах (на улицах, пешеходных переходах, вблизи светофоров). В случае использования видеозаписи она может быть преобразована в изображения путем дискретизации.

Анализ плотности транспортного потока – входные данные обрабатываются моделью анализа транспортной загруженности, при этом могут быть реализованы два подхода:

Обучение с учителем – данный метод требует наличия размеченных данных, что подразумевает разделение данных на входные и целевые выходные значения. Модель обучается в соответствии с желаемыми выходными данными. После обучения модель принимает изображения в качестве входных данных и выдает результат, представляющий собой оценку плотности транспортного потока. В данном случае модель извлекает признаки из данных и обучается находить зависимости между этими признаками и желаемым выходом [2].

Обучение без учителя – этот метод использует неразмеченные данные. Модель функционирует как система распознавания образов, выявляя движущиеся объекты на дороге, такие как автомобили, грузовики, велосипеды и другие транспортные средства. Для принятия решения о плотности потока модель использует пороговое значение. Принцип работы следующий:

Если количество распознанных объектов превышает пороговое значение → принимается решение о высокой загруженности дороги.

Выходные данные – это решение о состоянии дорожного движения. Полученное решение может быть передано аппаратным средствам, например, для переключения сигналов светофоров, либо отправлено в виде уведомлений водителям, чтобы они могли скорректировать свой маршрут в соответствии с текущей дорожной обстановкой [2].

Некоторые исследователи используют Vision Transformers (VT) в сочетании с сверточными нейронными сетями (CNN) для прогнозирования транспортных заторов в городских условиях в масштабах всего города. В предлагаемой архитектуре изображение дорожного движения подается на CNN, которая генерирует карты признаков. Затем эти карты признаков передаются в VT, где применяются две ключевые техники: токенизация и проекция. Токенизация используется для преобразования признаков в токены, содержащие информацию о зрительном восприятии, которые затем отправляются в проекцию, где трансформируются обратно в карты признаков и, в конечном итоге, передаются в LSTM [9]. Экспериментальные результаты показывают, что метод прогнозирования с использованием Vision Transformer на основе пространственно-временных характеристик является эффективным способом предсказания транспортного потока, особенно в условиях аномальных дорожных ситуаций. Предлагаемая технология превосходит традиционные методы по таким параметрам, как точность, полнота и воспроизводимость, а также способствует энергосбережению.

Структура модели глубокого обучения может быть унаследована от предварительно обученных моделей, таких как Yolo, Resnet или любой другой модели, но этого недостаточно, для этого требуются другие операции [Ошибка! Источник ссылки не найден.]:

Операции тонкой настройки путем обучения на основе собранных данных.

Оптимизация гиперпараметров: этот этап будет очень важен для наилучшего выбора параметров модели, таких как скорость обучения, количество слоев, размер пакета, размер ядра и т.д.

Адаптивное использование модели, это означает, что модель не будет одинаковой для всех областей или камер.

Структура обычно содержит слои свертки, слои объединения, слои нормализации, слои ослабления, рекуррентные слои, слои долговременной и кратковременной памяти.

#### **Список литературы:**

1. Батенов Ф. К. Искусственный интеллект в области дорожного движения //Международный форум молодых ученых. – 2021. – С. 22-24.
2. Галимов Р. Г. Основы алгоритмов машинного обучения-обучение с учителем //Аллея науки. – 2017. – Т. 1. – №. 14. – С. 810-817.
3. Скрипачев В. О. и др. Исследование сверточных нейронных сетей для обнаружения объектов на аэрокосмических снимках //International Journal of Open Information Technologies. – 2022. – Т. 10. – №. 7. – С. 54-64.
4. Хуссейн А. Х. С., Заргарян Е. В., Заргарян Ю. А. Модель прогнозирования транспортного потока на основе нейронных сетей для предсказания трафика на дорогах //Известия Южного федерального университета. Технические науки. – 2021. – №. 6 (223). – С. 124-132.
5. Ярошенко И. К. Внедрение искусственного интеллекта в государственное управление // «Эффективное государственное и муниципальное управление как фактор социально-экономического развития территорий», IV Всероссийская научно. – С. 152-154.
6. Akhatov A., Nazarov F., Eshtemirov B. Detection and analysis of traffic jams using computer vision technologies //Artificial Intelligence, Blockchain, Computing and Security Volume 2. – CRC Press, 2023. – С. 761-765.
7. Azfar T. et al. Deep learning-based computer vision methods for complex traffic environments perception: A review //Data Science for Transportation. – 2024. – Т. 6. – №. 1. – С. 1.
8. Basheer Ahmed M. I. et al. A real-time computer vision-based approach to detection and classification of traffic incidents //Big data and cognitive computing. – 2023. – Т. 7. – №. 1. – С. 22.
9. Ramana K. et al. A vision transformer approach for traffic congestion prediction in urban areas // IEEE Transactions on Intelligent Transportation Systems. – 2023. – Т. 24. – №. 4. – С. 3922-3934.

#### **References:**

1. Batenov F.K. Artificial Intelligence in the Field of Road Traffic // International Forum of Young Scientists. – 2021. – P. 22-24.
2. Galimov R.G. Fundamentals of Machine Learning Algorithms - Supervised Learning // Alley of Science. – 2017. – Vol. 1. – No. 14. – P. 810-817.

3. Skripachev V.O. et al. Study of Convolutional Neural Networks for Detecting Objects in Aerospace Images // International Journal of Open Information Technologies. - 2022. - Vol. 10. - No. 7. - P. 54-64.
4. Hussein A. H. S., Zargaryan E. V., Zargaryan Yu. A. A model for predicting traffic flow based on neural networks for predicting road traffic // Bulletin of the Southern Federal University. Technical sciences. - 2021. - No. 6 (223). - P. 124-132.
5. Yaroshenko I. K. Implementation of artificial intelligence in public administration // "Efficient state and municipal administration as a factor in the socio-economic development of territories", IV All-Russian scientific. - P. 152-154.
6. Akhatov A., Nazarov F., Eshtemirov B. Detection and analysis of traffic jams using computer vision technologies // Artificial Intelligence, Blockchain, Computing and Security Volume 2. - CRC Press, 2023. - pp. 761-765.
7. Azfar T. et al. Deep learning-based computer vision methods for complex traffic environments perception: A review // Data Science for Transportation. - 2024. - T. 6. - No. 1. - P. 1.
8. Basheer Ahmed M. I. et al. A real-time computer vision-based approach to detection and classification of traffic incidents // Big data and cognitive computing. - 2023. - T. 7. - No. 1. - P. 22.
9. Ramana K. et al. A vision transformer approach for traffic congestion prediction in urban areas // IEEE Transactions on Intelligent Transportation Systems. - 2023. - T. 24. - No. 4. - pp. 3922-3934.