

Научная статья  
УДК 629.7.05

## Актуальные вопросы совершенствования технического зрения при использовании на автомобилях

Александр Георгиевич Тыняный<sup>1</sup>, Сергей Робертович Кристальный<sup>2</sup>,  
Павел Александрович Красавин<sup>3</sup>, Максим Алексеевич Топорков<sup>4</sup>,  
Алексей Николаевич Андреев<sup>5</sup>

<sup>1,2,3,4,5</sup>Московский автомобильно-дорожный государственный технический университет (МАДИ),  
Москва, Россия.

<sup>1</sup>a.tynyanuy@madi.ru

<sup>2</sup>sporauto@mail.ru

<sup>3</sup>krasavin.madi@yandex.ru

<sup>4</sup>makstoporkov@rambler.ru

<sup>5</sup>andr\_aleksei@mail.ru

**Аннотация.** Системы технического зрения активно используются на автомобильном транспорте. Развитие этого направления приводит к увеличению числа бортовых камер, установленных на транспортном средстве. При помощи различных алгоритмов анализа изображений, с применением нейронных сетей можно получать информацию об окружающем пространстве автомобиля и свойствах рядом расположенных объектов. Важной особенностью является позиционирование объектов в пространстве и определение расстояния до них. Такие функции используются для определения безопасной дистанции и с целью построения оптимальной траектории движения автомобиля в заданном пространстве. В данной статье рассматриваются различные способы анализа изображений с целью применения на автомобилях в системах помощи водителю или в беспилотных автомобилях.

**Ключевые слова:** техническое зрение, обнаружение окружающего пространства, определение расстояния по изображению, нейронная сеть, системы помощи водителю.

**Для цитирования:** Тыняный А.Г., Кристальный С.Р., Красавин П.А. Топорков М.А., Андреев А.Н. Актуальные вопросы совершенствования технического зрения при использовании на автомобилях // Автомобиль. Дорога. Инфраструктура. 2026. № 1 (47).

Original article

## Current issues in improving technical vision for use in vehicles

Aleksandr G. Tynyanyy<sup>1</sup>, Sergey R. Kristalnyi<sup>2</sup>, Pavel A. Krasavin<sup>3</sup>, Maksim A. Toporkov<sup>4</sup>, Aleksey N. Andreev<sup>5</sup>

<sup>1,2,3,4,5</sup>Moscow Automobile and Road Construction State Technical University (MADI), Moscow, Russia

<sup>1</sup>a.tynyanyy@madi.ru

<sup>2</sup>sporauto@mail.ru

<sup>3</sup>krasavin.madi@yandex.ru

<sup>4</sup>makstoporkov@rambler.ru

<sup>5</sup>andr\_aleksei@mail.ru

**Abstract.** Technical vision systems are actively used in automotive transport. The development of this area leads to an increase in the number of onboard cameras installed on vehicles. Using various image analysis algorithms and neural networks, it is possible to obtain information about the vehicle's surroundings and the properties of nearby objects. An important feature is the positioning of objects in space and the determination of their distance. These functions are used to determine a safe distance and to create an optimal vehicle trajectory in a given space. This article discusses various methods of image analysis for use in car assistance systems or even in self-driving cars.

**Keywords:** technical vision, environmental detection, image-based distance measurement, neural networks, and driver assistance systems.

**For citation:** Tynyanyy A.G., Kristalnyi S.R., Krasavin P.A., Toporkov M.A., Andreev A.N. Current issues in improving technical vision for use in vehicles. *Avtomobil'. Doroga. Infrastruktura.* 2026. № 1 (47).

### Введение

Анализ существующих способов снижения числа дорожных происшествий на дорогах показывает, что одним из основных направлений является повышение безопасности со стороны конструкции транспортного средства и, в частности, применение систем ADAS (Advanced Driver Assistance Systems – интеллектуальные системы помощи водителю). Возникает задача в создании новых функций систем и увеличении эффективности их действия.

Работа таких систем осуществляется за счет оснащения автомобиля различными датчиками и сенсорами. Тенденция последних лет заключается в усовершенствовании систем кругового обзора транспортных средств. Современные автомобили все чаще оснащаются различными средствами для обнаружения окружающего пространства, такими как камеры, радары или лидары. Широкое распространение в конструкции автомобиля получают камеры в силу их низкой стоимости и широкого диапазона для применения. Помимо использования самих камер производители внедряют различные алгоритмы работы систем обнаружения окружающего пространства для дальнейшего функционирования систем безопасности автомобиля. К основным решениям в вопросах обработки изображений являются классификация и сегментация.

### **Выделение объектов на изображении через системы технического зрения**

Обнаружение объектов заключается в определении классов объектов, представленных на изображении. Количество обнаруживаемых классов объектов зависит от назначения системы и ограничен по причине необходимости высокой скорости работы. К примеру, в системе технического зрения автомобиля классами обнаруживаемых объектов могут являться автомобили, пешеходы, светофоры, линии разметки и т.д. Процесс обнаружения объектов должен выдавать набор аннотаций, содержащий метки классов объектов, а также набор оценок достоверности и ограничивающих рамок, которые охватывают объекты на изображении.

В задачах классификации изображений каждый объект маркируется соответствующей меткой. При наличии нескольких объектов на изображении, для каждого из них устанавливается своя собственная метка. Комплексное решение задач классификации и локализации предполагает определение местоположения единственного экземпляра объекта в кадре, даже в ситуациях, когда объектов несколько.

Другим типом определения объектов является сегментация. Сегментация изображений – это процесс разделения изображения на различные части или объекты для последующей обработки, анализа и классификации. В отличие от обычной классификации изображений, сегментация позволяет получить более глубокое понимание объектов и их пространственных взаимосвязей. Существует несколько типов сегментации, которые различаются по методу и целям [1].

*Двойная сегментация* – наиболее простой тип сегментации, который подразумевает разделение изображения на два класса: объект и фон. Главная цель – обособление определенного объекта на картинке. Данный метод часто применяется в простых задачах, например таких как отделение текста от фона или для идентификации границ предметов для обнаружения автомобиля на дороге.

В *мультиклассовой сегментации* к каждому пикселю изображения присваивается один из нескольких классов, которые являются взаимно исключающими. Актуально в задачах, где необходимо идентифицировать различные типы объектов в одном кадре. Например, при распознавании транспортных средств, зданий, проезжей части и пешеходов на фотографиях городских улиц.

*Семантическая сегментация* – это классификация каждого пикселя изображения в соответствии с заранее установленными классами. Целью является классифицировать каждый элемент изображения, относящегося к определенной категории (например, дороги, деревья, машины и т.п.). Используется для распознавания объектов на уровне категорий. Например, определение зон на изображении, занятых автомобилями, строениями или людьми (рис. 1). Нужно отметить, что семантическая сегментация не делает различий между отдельными экземплярами одного класса. Для определения количества объектов применяется *инстанс-сегментация* (*сегментация экземпляров*).



Рис. 1. Пример семантической сегментации изображения [2]

*Паноптическая сегментация* представляет собой интеграцию семантической и инстансной сегментации, в рамках которой каждому пикселю изображения присваивается либо идентификатор экземпляра, либо категория класса (семантическая основа). Основная задача заключается в создании объединенной системы сегментации, требующей от модели распознавания как отдельных объектов (например, транспортных средств, прохожих), так и общего фона (например, дорожного покрытия, архитектурных сооружений). Данный метод находит применение в сложных сценариях, где необходимо принимать во внимание как объекты, так и окружающую среду. К примеру изображения, на которых осуществляется классификация и разделение всех объектов (людей, автомобилей), а также элементов заднего плана (улиц, зданий, неба). Характерной чертой паноптической сегментации является то, что каждый пиксель должен быть соотнесен либо с определенным объектом, либо с классом, представляющим фоновое окружение.

### **Принципы анализа изображений с применением нейронных сетей**

В современных автомобильных системах компьютерного зрения широко используются алгоритмы, основанные на нейросетях с различными

структурами. Например, в модулях управления автомобилями Tesla Motors применяются многозадачные нейронные сети, объединенные в архитектуру HydraNet, являющуюся уникальной разработкой данной компании [3]. Эта сеть выполняет широкий спектр задач, связанных с анализом дорожной ситуации, включая идентификацию окружающих объектов и определение их характеристик, распознавание и толкование дорожных знаков, изучение дорожного покрытия и т.д.

Внедрение алгоритмов нейронных сетей в компьютерное зрение объясняется их значительным превосходством в области анализа изображений по сравнению с традиционными алгоритмами, созданными до появления нейросетей [4]. В основе этих алгоритмов лежат сверточные нейронные сети (CNN), представляющие собой математические модели биологических зрительных систем. Принцип работы нейросети исключает использование математических формул для описания анализируемых или прогнозируемых объектов и процессов.

Для того, чтобы нейронная сеть была работоспособна ее необходимо обучить с использованием предварительно собранных данных об объекте с массивами его входных и выходных переменных в различных рабочих условиях. В ходе обучения внутренние параметры сети настраиваются таким образом, чтобы она как можно точнее воспроизводила отклик выходных переменных объекта на заданные входные сигналы. Используемые в системах технического зрения алгоритмы глубокого обучения нейронных сетей (англ. Deep learning) основаны на оптимальном подборе параметров с использованием метода стохастического градиентного спуска [5, 6].

Подходы, основанные на нейронных сетях для идентификации объектов, разделяются на две основные категории. К первой относятся системы, реализующие двухшаговый процесс обнаружения. Примером этой группы является архитектура Faster R-CNN (Region Based Convolutional Neural

Network) [7]. Данные алгоритмы первоначально выявляют на изображении вероятные зоны расположения объектов, используя сеть, генерирующую региональные предложения RPN (Region Proposal Network) [8]. Затем, эти отобранные зоны передаются в сверточную нейронную сеть, базирующуюся на регионах R-CNN [9]. Эта нейросеть производит оценку вероятности принадлежности обнаруженного объекта в обозначенной зоне к конкретной категории, параллельно определяя его координаты на изображении. Ко второй категории относятся одноэтапные системы, такие как YOLO, SSD и RetinaNet. Подобные алгоритмы, как правило, базируются на полностью сверточных сетях и предоставляют информацию о классовой принадлежности объекта и его пространственном положении непосредственно.

В то время как двухэтапные детекторы обеспечивают превосходную точность в определении местоположения и идентификации объектов, одноэтапные детекторы, напротив, сразу предоставляют данные об объектах на изображении, минуя этап выделения регионов, потенциально содержащих объект. Следовательно, последние характеризуются большей эффективностью в плане скорости обработки и пригодны для использования в системах, работающих в режиме реального времени. Рассмотрим структуру нескольких сверточных сетей подробнее.

Сеть U-Net представляет собой нейросетевую архитектуру для решения задач сегментации изображений, особенно в области медицинской визуализации [1]. Главной отличительной чертой U-Net является её симметричная структура, напоминающая U-образную форму, в которой на этапе кодирования свёрточные слои постепенно уменьшают размер изображения, а на этапе декодирования происходит восстановление исходного разрешения (рис. 2). Такая схема позволяет сети эффективно сохранять пространственную информацию.

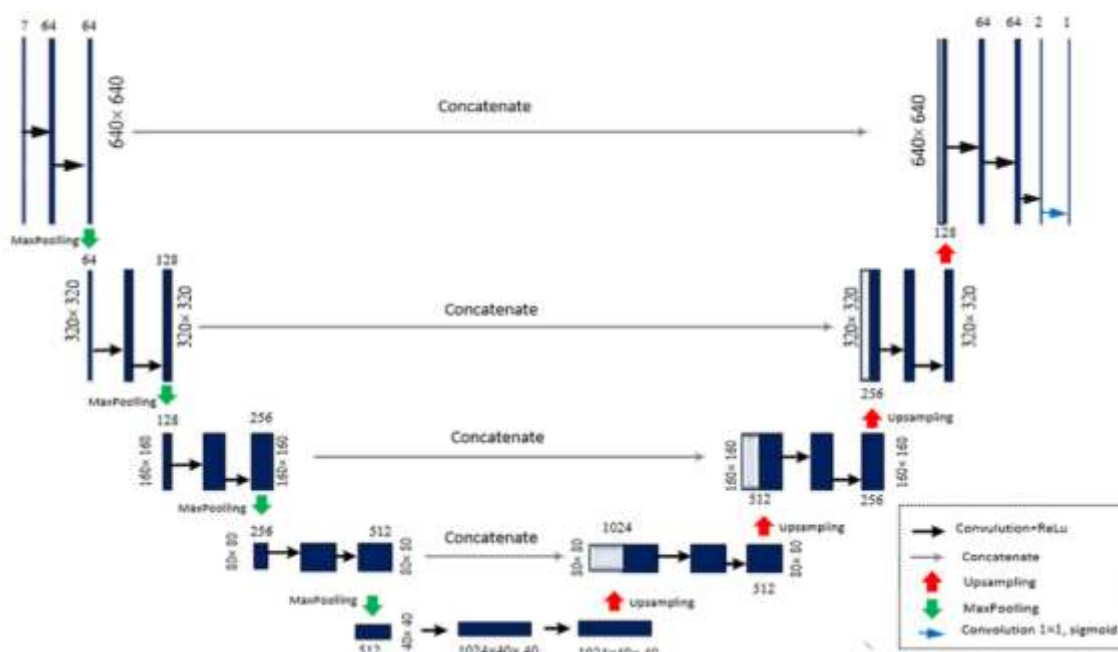


Рис. 2. Архитектура сети U-Net

Основная структура U-Net состоит из энкодера (Contracting Path), декодера (Expanding Path) и Скип-соединений (Skip Connections).

Энкодер, представляющий собой сверточную нейронную сеть, отвечает за извлечение характеристик из входного изображения и уменьшение его пространственного разрешения за счет применения сверточных операций и пулинга. Каждый блок энкодера состоит из двух последовательных сверточных слоев с функцией активации, за которыми следует Max Pooling, предназначенный для уменьшения размерности. Число фильтров возрастает с каждым уровнем. На данном этапе происходит сжатие входных данных и выделение наиболее значимых признаков.

Декодер восстанавливает пространственное разрешение, используя транспонированные свертки. Каждый блок декодера включает в себя транспонированную свертку и две свертки с функцией активации. Главная задача декодера – возвращение исходного размера изображения и классификация каждого пикселя.

В U-Net используются обходные соединения (skip connections), которые соединяют соответствующие слои энкодера и декодера, сохраняя

информацию, полученную на начальных этапах свертки, что крайне важно для достижения высокой точности локализации объектов на изображении.

Недостатками сети U-Net являются высокие вычислительные затраты и сложность восстановления очень мелких деталей или объектов, перекрытых на изображении другими.

Архитектура *FCN (Fully Convolutional Networks)* со своим появлением стала прогрессивной в области семантической сегментации изображений. Отличительной чертой FCN является замена полносвязных слоев, типичных для сетей классификации изображений, на сверточные слои. Это позволяет обрабатывать изображения произвольного размера и формировать карты сегментации для всей области просмотра (рис. 3).

Детализированная сегментация может быть затруднена из-за потери информации при операциях пулинга и свертки, несмотря на наличие обходных соединений. Кроме того, для сложных сцен с большим количеством объектов и классов FCN может оказаться недостаточно производительной, в таких случаях предпочтительнее использовать архитектуры типа U-Net или DeepLab.

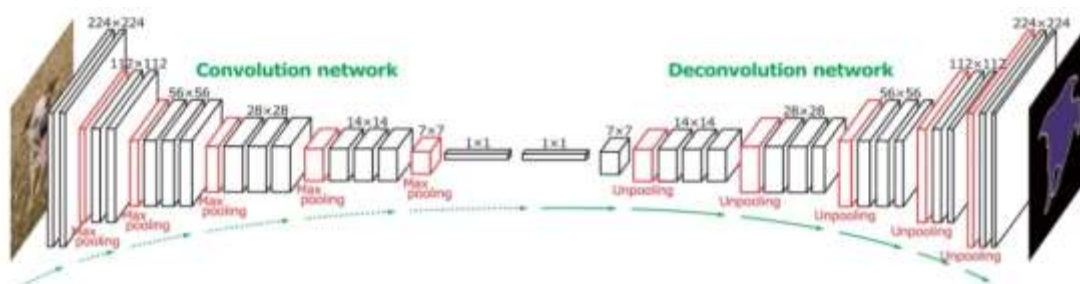


Рис. 3. Архитектура сети FCN

Сеть *SegNet* представляет собой сверточную нейронную сеть, предназначенную для семантической сегментации изображений. Ее ключевая особенность – эффективное решение задач сегментации с умеренными вычислительными затратами, что особенно важно для приложений, требующих высокого разрешения.

Для восстановления пространственного разрешения в SegNet используются индексы пула из энкодера, что повышает качество сегментации и снижает вычислительную нагрузку. Архитектура SegNet состоит из энкодера, декодера и классификационного слоя. Энкодер извлекает признаки, используя сверточные слои и операции Max Pooling, при этом сохраняя индексы пула. Декодер восстанавливает исходное разрешение изображения, используя эти индексы. Классификационный слой присваивает каждому пикселю метку класса (рис. 4).

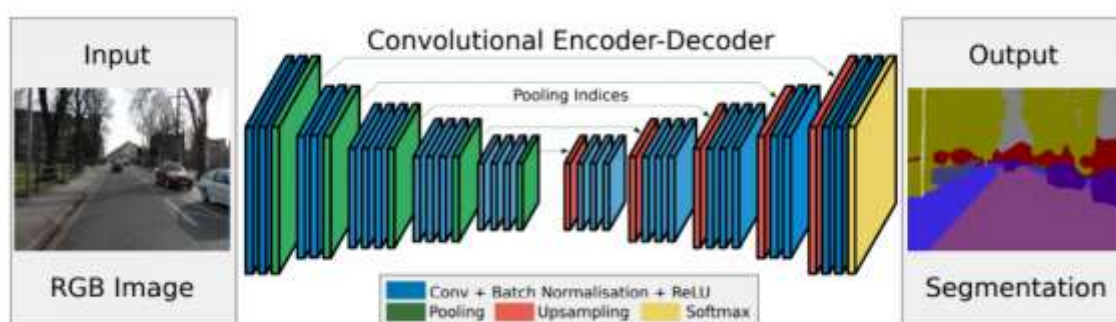


Рис. 4. Архитектура сети SegNet

Несмотря на свою эффективность, SegNet может уступать более сложным моделям, таким как U-Net или DeepLab, особенно на сложных сценах.

*DeepLab* – это группа нейросетевых архитектур, разработанных в Google DeepMind для решения задач семантической сегментации изображений. Основная цель DeepLab заключается в повышении точности сегментации объектов путём определения классов каждого пикселя с использованием продвинутых техник, таких как пространственное пирамидальное расширенное свертывание (ASPP) и модули декодирования для восстановления пространственного разрешения.

DeepLab, начиная от первой версии и до версии DeepLabv3+, продемонстрировала значительный прогресс в задачах сегментации, включая выделение объектов, обработку сцен в реальном времени и других задачах компьютерного зрения.

В первой версии DeepLabv1 использовались свёрточные нейронные сети, адаптированные для сегментации с применением атриусных свёрток (Atrous Convolutions) вместо стандартных. Атриусные свёртки позволяют расширить поле зрения без уменьшения разрешения карты признаков, что важно для обнаружения объектов разных размеров. Применяется модель CRF (Conditional Random Field) для постобработки результатов сегментации, улучшая границы объектов.

Во второй версии DeepLabv2 предложена концепция Atrous Spatial Pyramid Pooling (ASPP), использующая несколько параллельных атриусных свёрток. ASPP применяет параллельные свёртки с разными коэффициентами расширения, захватывая информацию на разных масштабах. Pre-trained networks: использование обученной сети ResNet для извлечения признаков.

DeepLabv3 улучшает ASPP, добавляя уровень свёрток и глобальное усреднение. В этой версии больше не используется CRF из-за улучшения ASPP и сети извлечения признаков.

В последней версии DeepLabv3+ происходит комбинация DeepLabv3 и декодирующего механизма, подобного U-Net. Применение декодера способствует более качественному возврату к исходному разрешению визуальных данных по завершении этапа кодирования. DeepLabv3+ сохраняет применение усовершенствованного ASPP, как и в предшествующей версии. В добавок применяется глубокая нейронная сеть для извлечения характерных черт (ResNet-101) в роли энкодера позволяет добиться повышения эффективности распознавания признаков.

По завершении задач сегментации определенные объекты возможно интегрировать в пространственную модель восприятия окружающей среды для оценки его траектории и дистанции до него.

### **Оценка расстояний до объектов на изображении**

Стандартные методы обнаружения объектов обычно ограничиваются идентификацией категории объекта и его позиции на картинке, игнорируя

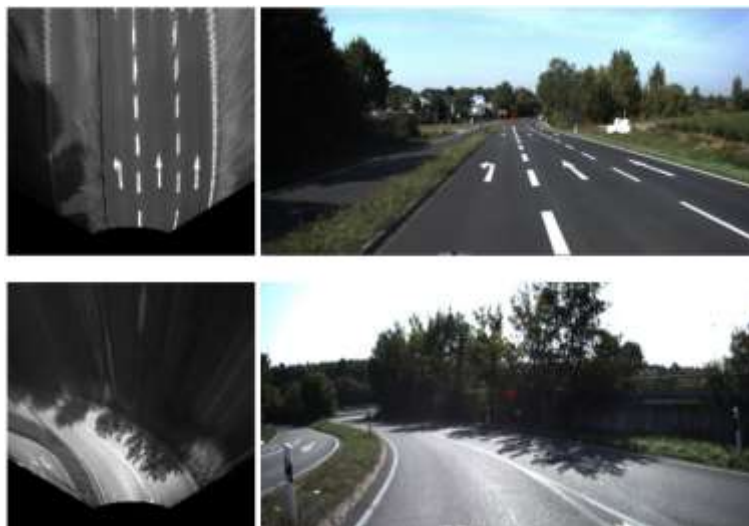
другие детали. Определение дистанции до объекта обычно основывается на информации, полученной от систем обнаружения (класс и параметры ограничивающего прямоугольника). На этой основе происходит вычисление удаленности объекта.

Основная идея большинства исследований в области оценки расстояния до окружающих объектов заключается в создании моделей, отображающих геометрическое соответствие между пикселями на изображении и их реальными физическими характеристиками в трехмерном пространстве. Распространенным методом оценки расстояния до объекта, выделенного точкой или ограничивающей рамкой, является преобразование координат пикселей в координаты топографически привязанного спутникового снимка с помощью алгоритма обратного перспективного отображения (Inverse Perspective Mapping, IPM) [10].

При распознавании дорожной разметки на изображениях, полученных с автомобильной камеры, перспективные искажения обуславливают изменение размеров разметки в зависимости от удаленности от камеры. Следовательно, для адекватного определения разметки требуется сопоставление с изменяющимися шаблонами, соответствующими ее местоположению на снимке. Перспектива придает пикселям разную значимость, зависящую от их координат.

После корректировки, учитывающей эффект перспективы, каждый пиксель начинает соответствовать участку дорожного полотна одинаковой площади, что обеспечивает равномерное распределение информации по всему изображению. Для компенсации влияния перспективы необходимо располагать информацией о параметрах камеры (положение, ориентация, фокусное расстояние), а также о геометрии запечатленной сцены (дорога, рассматриваемая как ровная поверхность).

Если рассматривать пример с дорожной разметкой (рис. 5), то становится ясно, что теперь ее характеристики не зависят от координат пикселя и сохраняют постоянство в любой точке изображения.



*Рис. 5. Пример обратной перспективы на изображениях автомобильной дороге [10]*

Следует учитывать, что ИРМ наиболее эффективен вблизи автомобиля. По мере увеличения расстояния геометрические характеристики объектов могут искажаться из-за неравномерного отображения. Это обстоятельство ограничивает дальность действия приложений и накладывает ограничения на точность результатов на больших расстояниях.

Для корректной работы ИРМ необходима калибровка камеры, определяющая ее внутренние и внешние параметры. Важно отметить, что стабильность этих параметров критически важна – их изменение при работе с разными типами дорожных покрытий снижает эффективность метода. В частности, прямое применение ИРМ, настроенного для одной дорожной поверхности, к другой, имеющей иной рельеф, может привести к искажению перспективы.

Для вычисления дистанций возможно использование сверточных нейронных сетей. Архитектура типа DisNet [11] служит примером подобного подхода. Как и методы обратного преобразования перспективы, этот способ основывается на принципах геометрии. DisNet представляет собой метод

определения дистанции до объектов, применяя для этого входящее изображение с единственной камеры и нейронную сеть. Первоначально, изображение, поступающее с камеры, обрабатывается классификатором объектов, основанным на алгоритме YOLO (You Only Look Once). Итогом работы являются границы обнаруженных объектов с идентификацией их классов. Далее, эти границы анализируются с целью извлечения характерных черт и параметров описывающих прямоугольников. Опираясь на эти входные параметры, обученная нейросеть DisNet производит оценку приблизительного расстояния между объектом и камерой.

Основной принцип работы метода заключается в использовании связи между настоящим размером объекта определенного типа и тем, как он выглядит на картинке. К примеру, для типа «человек» стандартные параметры, такие как рост, ширина и глубина, задаются заранее и известны для каждого распознаваемого типа.

Имея сведения о дистанциях до окружающих объектов, возникает вопрос о нахождении наилучшего или безопасного пути следования. Алгоритмы прокладки маршрута, задействующие методы искусственного интеллекта, решают задачу генерации траектории, основываясь на «обучении» с использованием примеров действий человека либо посредством системы штрафов и вознаграждений. Существуют разнообразные методы, позволяющие сформировать траекторию движения транспортного средства в классическом понимании [12], или даже предсказывать последовательность углов поворота руля, опираясь на данные с камеры машинного зрения, установленной в передней части автомобиля [13].

Нейросети оказались полезными и в области планирования маршрутов для роботизированного транспорта и беспилотных машин. Исследования показывают, что интеграция нейронных сетей с такими методами, как нечеткая логика или генетические алгоритмы, позволяет улучшить стабильность, ускорить процесс обучения нейросети и повысить общую

производительность алгоритма. Например, в исследовании [14] для навигации применяется нейронечеткий регулятор, где логика нечеткого вывода строится на основе заранее определенных базовых манёвров. Нейронная сеть включается в работу при активации примитива поворота и рассчитывает угол поворота руля, необходимый для выполнения этого маневра.

Альтернативный подход – сквозное обучение [15], при котором нейронные сети напрямую преобразуют визуальную информацию о дорожной обстановке в управляющие команды (рис. 6).

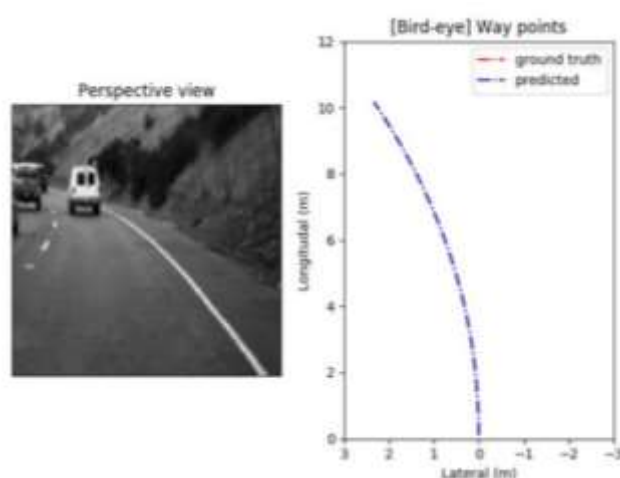


Рис. 6. Определение курса движения внутри полосы с использованием нейронной сети

На рисунке 6 слева представлен кадр, зафиксированный автомобильной камерой, закрепленной на капоте. Справа отображена траектория движения в декартовых координатах, сформированная на базе анализа полученного изображения с помощью нейросети.

Одним из многообещающих направлений использования нейронных сетей является обучение с подкреплением (Reinforcement Learning, RL). Данный метод позволяет объекту, выполняющему поиск оптимальной траектории, приспосабливать свою динамику движения к окружающей обстановке, невзирая на размеры территории или число препятствий вокруг транспортного средства.

## Заключение

Таким образом, существует множество применений алгоритмов не только в определении классов объектов и окружающего пространства автомобиля, но и определении расстояний до этих объектов. Также дальнейшим развитием технического зрения служат нейросетевые алгоритмы в области определения оптимальной траектории движения по ограниченному пространству. Такие ограничения могут быть заданы координатным способом, или линиями дорожной разметки.

Усовершенствование существующих алгоритмов или их совместное сочетание позволит повысить эффективность работы систем помощи водителю, что в свою очередь снизит риски возникновения ДТП, увеличит среднюю скорость транспортного средства на заданном маршруте.

## Список источников

1. Эволюция архитектур нейросетей в компьютерном зрении: сегментация изображений. – URL: <https://habr.com/ru/companies/sisoft/articles/864994/?ysclid=mi7e2qugg9143782482> (дата обращения: 24.10.2025).
2. Semantic Segmentation Using Deep Learning. – MATLAB & Simulink. – URL: <https://www.mathworks.com/help/vision/ug/semantic-segmentation-using-deep-learning.html#d119e211> (дата обращения: 20.12.2024).
3. Karpathy, A. PyTorch at Tesla / A. Karpathy // Научно-технический доклад и презентация, PyTorch DEVCON-19, 6 ноября 2019 г. – URL: [https://www.youtube.com/watch?v=oBklltKXtDE&ab\\_channel=PyTorch](https://www.youtube.com/watch?v=oBklltKXtDE&ab_channel=PyTorch) (дата обращения: 15.07.2025).
4. Goodfellow, I. Deep Learning / I. Goodfellow, Y. Bengio, A. Courville. – Massachusetts: MIT Press, 2016. – 800 p. – ISBN 0262035618.
5. Kingma, Diederik P. Adam: A Method for Stochastic Optimization / Diederik P. Kingma, J. Ba. – URL: <https://arxiv.org/abs/1412.6980> (дата обращения: 15.07.2025).
6. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks / S. Ren, K. He, R. Girshick, J. Sun // IEEE transactions on pattern analysis and machine intelligence. – 2017. – Vol. 39, No. 6. – P. 1137-1149. – DOI 10.1109/TPAMI.2016.2577031.

7. Region Proposal by Guided Anchoring / J. Wang, K. Chen, Sh. Yang, Ch.Ch. Loy // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. – 2019. – P. 2965-2974. – DOI 10.1109/cvpr.2019.00308.
8. Region-Based Convolutional Networks for Accurate Object Detection and Segmentation / R. Girshick, J. Donahue, T. Darrell, J. Malik // IEEE transactions on pattern analysis and machine intelligence. – 2016. – Vol. 38, No. 1. – P. 142-158. – DOI 10.1109/TPAMI.2015.2437384.
9. Семантическая сегментация с высоты птичьего полета для автономного вождения. – URL: <https://readmedium.com/monocular-birds-eye-view-semantic-segmentation-for-autonomous-driving-ee2f771afb59> (дата обращения: 05.10.2025).
10. Обратное преобразование перспективы с помощью ручной точки схода. – URL: <https://github.com/osvaldlaszlo/inverse-perspective-mapping?tab=readme-ov-file> (дата обращения: 05.10.2025).
11. DisNet: a novel method for distance estimation from monocular camera / MA. Haseeb, J. Guan, D. Ristic-Durrant, A. Gräser // IEEE/RSJ international conference on intelligent robots and systems – IROS 2018, 10th workshop on planning, perception and navigation for intelligent vehicles (PPNIV). – 2018.
12. Nair, R.S. Robotic Path Planning Using Recurrent Neural Networks / R.S. Nair, P. Supriya // 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT). – IEEE, 2020. – P. 1-5. – DOI 10.1109/ICCCNT49239.2020.9225479.
13. Multimodal End-to-End Autonomous Driving / Yi. Xiao, F. Codevilla, A. Gurram [et al.] // IEEE Transactions on Intelligent Transportation Systems. – 2022. – Vol. 23, No. 1. – P. 537-547. – DOI 10.1109/tits.2020.3013234. – EDN AVEFER.
14. Behavior-based neuro-fuzzy controller for mobile robot navigation / P. Rusu, E.M. Petriu, T.E. Whalen, A. Cornell, H.J.W. Spoelder // IEEE Transactions on Instrumentation and Measurement. – 2003. – Vol. 52, No. 4. – P. 1335-1340. – DOI 10.1109/TIM.003.816846.
15. End-to-End Deep Neural Network Design for Short-term Path Planning / M. Dao, D. Lanza, V. Fremont. – URL: <https://hal.archives-ouvertes.fr/hal-02266802> (дата обращения: 15.07.2025).

### References

1. Evolyutsiya arkhitektur neyrosetey v komp'yuternom zrenii: segmentatsiya izobrazheniy, available at: <https://habr.com/ru/companies/sloft/articles/864994/?ysclid=mi7e2qugg9143782482> (24.10.2025).
2. Semantic Segmentation Using Deep Learning, MATLAB & Simulink, available at: <https://www.mathworks.com/help/vision/ug/semantic-segmentation-using-deep-learning.html#d119e211> (20.12.2024).

3. Karpathy A. PyTorch at Tesla, *available at:* [https://www.youtube.com/watch?v=oBklltKXtDE&ab\\_channel=PyTorch](https://www.youtube.com/watch?v=oBklltKXtDE&ab_channel=PyTorch) (15.07.2025).
4. Goodfellow I., Bengio Y., Courville A. *Deep Learning*, Goodfellow, Massachusetts, MIT Press, 2016, 800 p., ISBN 0262035618.
5. Kingma Diederik P., Ba J. Adam: A Method for Stochastic Optimization, *available at:* <https://arxiv.org/abs/1412.6980> (15.07.2025).
6. Ren S., HeK., Girshick R., Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, *IEEE transactions on pattern analysis and machine intelligence*, 2017, vol. 39, no. 6, pp. 1137-1149, doi 10.1109/TPAMI.2016.2577031.
7. Wang J., Chen K., Yang Sh., Loy Ch.Ch. Region Proposal by Guided Anchoring, *IEEE/CVF conference on computer vision and pattern recognition*, Proceedings of the conference, 2019, pp. 2965-2974, doi 10.1109/cvpr.2019.00308.
8. Girshick R., Donahue J., Darrell T., Malik J. Region-Based Convolutional Networks for Accurate Object Detection and Segmentation, *IEEE transactions on pattern analysis and machine intelligence*, 2016, vol. 38, no. 1, pp. 142-158, doi 10.1109/TPAMI.2015.2437384.
9. Семантическая сегментация с высоты птичьего помета для автономного вождения, *available at:* <https://readmedium.com/monocular-birds-eye-view-semantic-segmentation-for-autonomous-driving-ee2f771afb59> (05.10.2025).
10. Обратное преобразование перспективы с помощью ручной точки шхода, *available at:* <https://github.com/osvaldlaszlo/inverse-perspective-mapping?tab=readme-ov-file> (05.10.2025).
11. Haseeb MA., Guan J., Ristic-Durrant D., Gräser A. DisNet: a novel method for distance estimation from monocular camera, *IEEE/RSJ international conference on intelligent robots and systems – IROS 2018, 10th workshop on planning, perception and navigation for intelligent vehicles (PPNIV)*, 2018.
12. Nair R.S., Supriya P. Robotic Path Planning Using Recurrent Neural Networks, *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, IEEE, 2020, pp. 1-5, doi 10/1109/ICCCNT49239.2020.9225479.
13. Xiao Yi., Codevilla F., Gurram A., Urfalioglu O., Lopez A.M. Multimodal End-to-End Autonomous Driving, *IEEE Transactions on Intelligent Transportation Systems*, 2022, vol. 23, no. 1, pp. 537-547, doi 10.1109/tits.2020.3013234.
14. Rusu P., Petriu E.M., Whalen T.E., Cornell A., Spoelder H.J.W. Behavior-based neuro-fuzzy controller for mobile robot navigation, *IEEE Transactions on Instrumentation and Measurement*, 2003, vol. 52, no. 4, pp. 1335-1340, doi 10.1109/TIM.003.816846.
15. Dao M., Lanza D., Fremont V. End-to-End Deep Neural Network Design for Short-term Path Planning, *available at:* <https://hal.archives-ouvertes.fr/hal-02266802> (15.07.2025).

Рецензент: А.Е. Есаков, канд. техн. наук, доц., Московский политехнический университет

## *Информация об авторах*

**Тыняный Александр Георгиевич**, аспирант кафедры «Автомобили», МАДИ.

**Кристалный Сергей Робертович**, канд. техн. наук,  
доцент кафедры «Автомобили», МАДИ.

**Красавин Павел Александрович**, канд. техн. наук,  
доцент кафедры «Автомобили», МАДИ.

**Топорков Максим Алексеевич**, канд. техн. наук,  
доцент кафедры «Автомобили», МАДИ.

**Андреев Алексей Николаевич**, ст. преподаватель кафедры «Автомобили», МАДИ.

## *Information about the authors*

**Tynyanu Aleksandr G.**, postgraduate student of the Department of Automobiles, MADI.

**Kristalniy Sergey R.**, Candidate of Sciences (Technical),  
associate professor of the Department of Automobiles, MADI.

**Krasavin Pavel A.**, Candidate of Sciences (Technical),  
associate professor of the Department of Automobiles, MADI.

**Toporkov Maksim A.**, Candidate of Sciences (Technical),  
associate professor of the Department of Automobiles, MADI.

**Andreev Aleksey N.**, senior lecturer of the Department of Automobiles, MADI.

*Статья поступила в редакцию 12.12.2025; одобрена после рецензирования 21.01.2026; принята к публикации 19.03.2026.*

*The article was submitted 12.12.2025; approved after reviewing 21.01.2026; accepted for publication 19.03.2026.*