

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
КЫРГЫЗСКОЙ РЕСПУБЛИКИ
КЫРГЫЗСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ ИМ.
И.РАЗЗАКОВА
ИНСТИТУТ ЭЛЕКТРОНИКИ И ТЕЛЕКОММУНИКАЦИЙ
Кафедра «ИНФОРМАЦИОННЫЕ СИСТЕМЫ ТЕХНОЛОГИИ
ИМ. АКАДЕМИКА А. ЖАЙНАКОВА»

ОТЧЕТ

О ПРОХОЖДЕНИИ НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ ПРАКТИКИ

Направление подготовки 710200 Информационные системы и технологии
программа: Информационные системы и технологии в горном деле

Место прохождения практики Инфо энергетика кафе, КГТУ.

Сроки прохождения практики с 20.01 2025 г. по 14.03. 2025 г.

Выполнил (а): магистрант (ка) группы ИСТм1-23 Давлетбаев Актан
Тургунбекович

(Ф.И.О)

Руководитель практики: к.ф.-м.н., доцент Аманкулова Н. А
(ученая степень, ученое звание, Ф.И.О.)

Бишкек – 2025

Содержание	
Введение.....	2
ГЛАВА 1. ОСНОВЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА.....	6
1.1 Основные понятия об искусственном интеллекте	6
1.2 Алгоритмы машинного обучения	14
1.3 Алгоритмы глубокого обучения.....	17
1.4 Автоматическое обнаружение объектов.....	18
1.5 Применение автоматического обнаружения объектов.....	22
1.6 Принципы работы систем обнаружения	23
ГЛАВА 2. АНАЛИЗ МЕТОДОВ ОБНАРУЖЕНИЯ И КЛАССИФИКАЦИИ	25
2.1 Классификация объектов	25
2.2 Сравнение различных алгоритмов.....	31
2.3 Проблемы и вызовы в автоматическом обнаружении и классификации	37
2.4 Результаты и обсуждение.....	43
Заключение	51
Список литературы.....	52

Введение

Научно-исследовательская практика является важным этапом в подготовке магистранта, направленным на углубленное изучение выбранной темы диссертационного исследования, овладение методами научного анализа и экспериментальной работы. Основной целью данной практики является развитие навыков самостоятельного научного исследования, включающего изучение современных методов, проведение экспериментов и анализ полученных данных.

В рамках данной научно-исследовательской практики проводится исследование в области автоматического обнаружения и классификации объектов с использованием методов искусственного интеллекта. Современные технологии глубокого обучения позволяют значительно повысить точность и скорость обработки визуальной информации, что делает их востребованными в различных сферах, включая медицинскую диагностику, автономный транспорт, системы безопасности и дистанционный мониторинг [1].

Настоящее исследование направлено на сравнительный анализ различных алгоритмов обнаружения и классификации объектов, выявление их преимуществ и недостатков, а также поиск путей их оптимизации. Практическая часть работы включает разработку и тестирование модели на реальных данных, что позволит оценить её эффективность и применимость в различных условиях [3,7].

Автоматическое обнаружение и классификация объектов на основе искусственного интеллекта приобретает всё большее значение в современном мире [6]. С каждым годом растет объем данных, требующих анализа: с камер видеонаблюдения, дронов, медицинских устройств, спутников, беспилотных автомобилей и т.д. Человеческие возможности по обработке и анализу этих данных ограничены, в то время как современные ИИ-системы способны обрабатывать большие объемы информации в кратчайшие сроки [6].

Актуальность темы подтверждается рядом факторов:

Рост объема данных и необходимость их автоматизированного анализа. Видеонаблюдение, спутниковые снимки, медицинские изображения — все это требует автоматизации анализа, чтобы оперативно извлекать важную информацию и принимать решения на основе данных [12].

Применение в критических областях. Автоматические системы на основе ИИ применяются в таких сферах, как безопасность, где важно своевременно обнаружить потенциальные угрозы (например, подозрительных людей или объекты на кадрах видеонаблюдения), или в медицине, где автоматическая классификация может помочь в диагностике заболеваний на основе медицинских изображений.

Автономные системы. Беспилотные автомобили и роботы нуждаются в высокоточном обнаружении и классификации объектов для правильного функционирования [10]. Это касается как обеспечения безопасности, так и эффективности их работы. Например, для беспилотного автомобиля важно правильно классифицировать объекты на дороге (пешеходы, транспортные средства, дорожные знаки) для предотвращения аварий [6].

Повышение эффективности. Использование ИИ позволяет сократить затраты времени и ресурсов, которые требуются для анализа данных человеком. Это становится особенно важным в условиях ограниченных ресурсов, например, при обработке больших массивов данных в реальном времени.

Технологические достижения. Современные методы глубокого обучения, такие как свёрточные нейронные сети (CNN), трансформеры и другие архитектуры, демонстрируют высокую точность в задачах распознавания и классификации объектов [3,10]. Это делает тему исследования особенно актуальной, поскольку с каждым годом технологии совершенствуются, открывая новые возможности для их применения.

Перспективы дальнейшего развития. С каждым годом количество приложений, использующих ИИ для автоматической классификации объектов,

только увеличивается. Ожидается, что в ближайшие десятилетия ИИ станет основой для новых прорывных технологий в различных сферах, начиная от робототехники и заканчивая космическими исследованиями.

Таким образом, исследование методов автоматического обнаружения и классификации объектов на основе искусственного интеллекта является не только актуальным, но и необходимым для дальнейшего развития различных технологий, которые обеспечивают повышение безопасности, эффективности и комфорта в жизни человека.

Цель исследования

Основная цель данного исследования заключается в анализе и оценке современных методов автоматического обнаружения и классификации объектов на основе искусственного интеллекта. В рамках этой цели предполагается изучить существующие подходы, алгоритмы и технологии, используемые для решения данных задач, а также выявить перспективные направления их дальнейшего развития и применения в различных отраслях.

Задачи исследования

Для достижения указанной цели необходимо решить следующие задачи:

Анализ существующих методов и подходов к обнаружению и классификации объектов на основе ИИ:

Провести обзор и классификацию современных методов, используемых для автоматического обнаружения и классификации объектов, таких как свёрточные нейронные сети (CNN), рекуррентные нейронные сети (RNN), трансформеры и др.

Технологии автоматического обнаружения и классификации объектов применяются в самых различных областях:

Автономное вождение. В этой сфере нейронные сети используются для распознавания пешеходов, транспортных средств, дорожных знаков и других объектов на дороге. Одним из ключевых примеров является использование

технологий глубокого обучения в системах автопилота, разрабатываемых компаниями Tesla, Waymo и другими.

Медицинская диагностика. В медицине нейронные сети активно применяются для анализа медицинских изображений (например, рентгеновских снимков, КТ и МРТ), а также для классификации заболеваний, таких как рак [7]. Исследования показывают, что системы на основе ИИ могут достигать точности, сопоставимой с экспертами в определенных задачах, что подтверждается работами Esteva et al. (2017).

Видеонаблюдение и безопасность. Методы ИИ широко используются для обнаружения подозрительных объектов, распознавания лиц и анализа поведения в системах безопасности. Модели, такие как YOLO и SSD, применяются для мониторинга в реальном времени.

Анализ спутниковых данных. Нейронные сети используются для автоматической классификации объектов на спутниковых снимках, что находит применение в сельском хозяйстве, градостроительстве, мониторинге природных ресурсов.

Несмотря на значительные успехи в разработке и применении методов ИИ, существуют и определенные вызовы, с которыми сталкиваются исследователи:

Необходимость больших объемов данных. Для обучения глубоких нейронных сетей требуются большие наборы данных, которые не всегда доступны для специфических задач.

Объяснимость решений. Одной из ключевых проблем является проблема интерпретируемости решений ИИ-систем [12]. Многие современные модели функционируют как "черные ящики", и их решения сложно объяснить, что вызывает вопросы относительно доверия к таким системам в критически важных приложениях.

ГЛАВА 1. ОСНОВЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

1.1 Основные понятия об искусственном интеллекте

Искусственный интеллект (ИИ) представляет собой направление информатики, целью которого является создание систем, способных выполнять задачи, которые традиционно требуют человеческого интеллекта. Это могут быть такие задачи, как восприятие, логическое мышление, обучение, принятие решений и выполнение действий в условиях неопределённости. ИИ широко применяется в различных областях, включая здравоохранение, финансы, промышленность, транспорт и многие другие.

Определение искусственного интеллекта

Искусственный интеллект можно разделить на две основные категории:

Слабый ИИ (Narrow AI): системы, предназначенные для выполнения конкретных задач. Эти системы могут превосходить человека в узких областях (например, распознавание изображений, игра в шахматы), но не обладают общей способностью к интеллектуальному мышлению или обучению на основе нового опыта. Примеры слабого ИИ: голосовые помощники (Siri, Alexa), системы распознавания лиц, автоматические переводы текста.

Сильный ИИ (General AI): гипотетический вид ИИ, который может выполнять любые интеллектуальные задачи, с которыми справляется человек. Сильный ИИ способен к абстрактному мышлению, самосознанию и может обучаться на опыте так же, как это делает человек. На данный момент сильный ИИ остаётся объектом теоретических исследований и не реализован на практике.

В рамках ИИ выделяются несколько подходов и парадигм, которые используются для построения интеллектуальных систем:

Логические системы и экспертные системы:

Экспертные системы были одними из первых успешных приложений ИИ. Эти системы используют набор правил (если-то) для моделирования рассуждений эксперта в узкой области (например, диагностика заболеваний).

Логика первого порядка (FOL) и другие формальные логические системы предоставляют математический аппарат для моделирования рассуждений. Несмотря на свою мощь, такие системы имеют ограниченное применение в реальных задачах, где не все условия могут быть формализованы.

Эволюционные алгоритмы, такие как генетические алгоритмы, используют концепции из биологии, такие как отбор, мутации и скрещивание, для оптимизации решений. Эти методы применяются в тех случаях, когда точное решение задачи сложно найти напрямую, и нужно исследовать большое пространство возможностей.

Архитектуры нейронных сетей

Одним из наиболее значимых достижений в области ИИ является развитие нейронных сетей, которые моделируют работу биологических нейронов в мозге. Основные архитектуры нейронных сетей включают:

Полносвязные нейронные сети (Fully Connected Neural Networks, FCNN): это базовый тип нейронной сети, где каждый нейрон связан с каждым другим нейроном на предыдущем и последующем уровне. Эти сети хорошо подходят для обработки табличных данных и применяются в задачах регрессии и классификации.

Свёрточные нейронные сети (Convolutional Neural Networks, CNN): CNN созданы специально для обработки изображений и данных с пространственной структурой. Они используют слои свёрток для выделения признаков изображения на разных уровнях, что делает их крайне эффективными в задачах распознавания объектов и классификации изображений.

Рекуррентные нейронные сети (Recurrent Neural Networks, RNN): эти сети используются для работы с последовательными данными, такими как текст или

временные ряды. RNN запоминают информацию из предыдущих шагов, что делает их полезными в задачах, где требуется учитывать контекст, например, в машинном переводе и распознавании речи.

Трансформеры (Transformers): это современная архитектура нейронных сетей, которая использует механизм внимания для работы с длинными последовательностями данных. Трансформеры достигли больших успехов в задачах обработки естественного языка (NLP) и начинают активно применяться в задачах компьютерного зрения.

Практическое применение искусственного интеллекта

ИИ находит широкое применение в различных сферах:

Здравоохранение: анализ медицинских изображений, диагностика заболеваний, разработка персонализированных планов лечения.

Транспорт: автономные транспортные средства, системы навигации, оптимизация маршрутов.

Финансы: алгоритмическая торговля, анализ кредитных рисков, обнаружение мошенничества.

Индустрия развлечений: персонализация рекомендаций, создание контента, виртуальные ассистенты.

Этические и социальные аспекты ИИ

С развитием ИИ возникает множество вопросов, связанных с его использованием. Основные этические проблемы включают:

Прозрачность и объяснимость: многие современные ИИ-системы являются "чёрными ящиками", и их решения сложно объяснить, что вызывает опасения в критически важных приложениях, таких как медицина и право.

Конфиденциальность данных: системы ИИ требуют больших объемов данных для обучения, что ставит под угрозу конфиденциальность личной информации пользователей.

Рабочие места и автоматизация: автоматизация на основе ИИ может приводить к вытеснению рабочих мест, особенно в рутинных профессиях, что требует разработки новых подходов к социальной политике и подготовке кадров.

Безопасность ИИ: важно разрабатывать системы ИИ, которые работают надёжно и безопасно, особенно в критически важных областях, таких как транспорт и оборона.

Искусственный интеллект (ИИ) — это область информатики, которая занимается созданием машин или программ, способных выполнять задачи, требующие человеческого интеллекта. ИИ стремится имитировать человеческие когнитивные функции, такие как обучение, рассуждение, планирование, восприятие и адаптация к новым ситуациям.

В более узком смысле, ИИ может быть определён как способность систем анализировать окружающую среду, принимать решения и действовать для достижения поставленных целей. В основе большинства современных систем ИИ лежат алгоритмы машинного обучения, которые позволяют системам улучшать свои результаты на основе предыдущего опыта.

Искусственный интеллект можно разделить на несколько основных категорий в зависимости от уровня сложности и целей:

Слабый ИИ (Narrow AI):

Определение: Слабый ИИ, также известный как узкий ИИ, разрабатывается для решения конкретных задач и не обладает способностью к общему интеллектуальному мышлению.

Системы распознавания речи и лиц.

Рекомендательные системы (например, YouTube, Netflix).

Виртуальные ассистенты (Siri, Google Assistant).

Особенности: Эти системы выполняют свои задачи с высокой точностью, но ограничены заранее запрограммированными областями. Они не могут

выходить за рамки своих возможностей и обучаться новым задачам без дополнительной тренировки.

Сильный ИИ (General AI):

Определение: Сильный ИИ представляет собой гипотетическую систему, способную решать широкий спектр задач, как и человеческий интеллект. Такие системы могут учиться и адаптироваться к новым ситуациям, обладать самосознанием и принимать решения на основе понимания окружающего мира.

Пример применения: На данный момент сильный ИИ остаётся целью исследований, и его реальные реализации отсутствуют.

Особенности: Эта система должна обладать всеми когнитивными функциями человека и способностью решать задачи без необходимости их предварительного программирования.

Сверхразумный ИИ (Superintelligence):

Определение: Сверхразумный ИИ представляет собой теоретическую форму ИИ, которая значительно превосходит человеческий интеллект во всех аспектах, включая решение задач, творчество и социальные навыки.

Пример применения: Этот вид ИИ пока остаётся гипотетическим, и его возможное появление вызывает множество этических и философских вопросов.

Особенности: Сверхразумный ИИ может быть не только более умным, чем человек, но и развивать новые формы мышления, не доступные человеческому разуму.

Виды ИИ по методам обучения

Машинное обучение (Machine Learning):

Определение: Машинное обучение — это подмножество ИИ, в котором машины обучаются на основе данных без явного программирования всех возможных решений. Модели машинного обучения обучаются на наборах данных и используют их для построения предсказаний.

Пример применения: Классификация изображений, системы рекомендаций, предсказательная аналитика.

Типы машинного обучения:

Обучение с учителем: Обучение на размеченных данных (когда известны правильные ответы).

Обучение без учителя: Обучение на неразмеченных данных для поиска скрытых закономерностей.

Обучение с подкреплением: Агент взаимодействует с окружающей средой, получая обратную связь в виде наград и наказаний, и стремится максимизировать награды.

Глубокое обучение (Deep Learning):

Определение: Глубокое обучение — это подмножество машинного обучения, которое использует многослойные нейронные сети для анализа больших объемов данных и обучения сложным закономерностям. Такие сети имитируют работу человеческого мозга, что делает их крайне эффективными для задач компьютерного зрения, обработки текста и звука.

Пример применения: Распознавание речи, автономное вождение, диагностика заболеваний.

Особенности: Глубокие нейронные сети могут автоматически извлекать признаки из данных, что делает их особенно полезными для работы с необработанными данными, такими как изображения и видео.

Эволюционные алгоритмы (Evolutionary Algorithms):

Определение: Эволюционные алгоритмы основаны на биологических процессах эволюции, таких как отбор, мутация и скрещивание. Эти методы используются для поиска оптимальных решений в сложных пространствах решений.

Пример применения: Оптимизация сложных инженерных задач, генетические алгоритмы для поиска решений в играх или финансовых задачах.

Особенности: Эти алгоритмы эффективны для решения задач с большими и сложными пространствами решений, но могут требовать больших вычислительных ресурсов.

Виды ИИ по применению

Когнитивные системы:

Определение: Когнитивные системы предназначены для выполнения задач, связанных с когнитивными функциями человека, такими как обучение, память и восприятие. Эти системы могут анализировать информацию, обучаться на основе опыта и взаимодействовать с пользователем на естественном языке.

Пример применения: IBM Watson, системы автоматизированной диагностики, чат-боты.

Особенности: Когнитивные системы обладают способностью адаптироваться и совершенствоваться по мере накопления опыта.

Робототехника (Robotics):

Определение: Робототехника использует ИИ для создания машин, способных автономно взаимодействовать с окружающей средой и выполнять физические задачи. Это могут быть как промышленные роботы, так и роботы для повседневного использования, например, в домах или больницах.

Пример применения: Роботы на заводах, автономные дроны, роботы-помощники.

Особенности: Роботы на основе ИИ могут обучаться и улучшать свои способности на основе анализа данных сенсоров, окружающей среды и взаимодействий.

Обработка естественного языка (Natural Language Processing, NLP):

Определение: NLP — это область ИИ, которая занимается взаимодействием человека и компьютера на естественном языке. Эти системы способны анализировать, понимать и генерировать текст или речь, что позволяет

создавать приложения для перевода, поиска информации, общения и других задач.

Пример применения: Машинный перевод (Google Translate), виртуальные ассистенты (Alexa, Siri), чат-боты.

Особенности: Современные достижения в NLP, особенно с использованием трансформеров, значительно улучшили точность и возможности обработки текста и речи.

1.2 Алгоритмы машинного обучения

Машинное обучение (ML) и глубокое обучение (DL) представляют собой два ключевых направления в искусственном интеллекте, которые лежат в основе современных интеллектуальных систем. Эти методы позволяют компьютерам обучаться на данных и принимать решения без явного программирования.

Машинное обучение включает различные подходы и алгоритмы, которые можно разделить на три основные категории: обучение с учителем, обучение без учителя и обучение с подкреплением.

Обучение с учителем (Supervised Learning) предполагает использование размеченных данных, то есть набора примеров, где для каждого входного примера известен правильный результат (метка). Модель обучается на этих данных, чтобы предсказывать результаты для новых, ещё не известных примеров.

Линейная регрессия (Linear Regression):

Описание: Алгоритм, который используется для предсказания числового значения на основе входных данных. Модель строит линейную зависимость между входными признаками и целевой переменной.

Пример применения: Прогнозирование цен на жильё, предсказание финансовых показателей.

Логистическая регрессия (Logistic Regression):

Описание: Используется для классификации бинарных данных. Логистическая регрессия предсказывает вероятность того, что объект принадлежит к одному из двух классов, используя функцию сигмоида для вывода вероятности.

Пример применения: Обнаружение спама, диагностика заболеваний.

Деревья решений (Decision Trees):

Описание: Строят последовательность решений в виде дерева, где каждая ветвь представляет вопрос, а каждый лист — окончательный ответ или категорию. Алгоритм прост для понимания и интерпретации.

Пример применения: Классификация видов растений, принятие бизнес-решений.

Случайный лес (Random Forest):

Описание: Это ансамблевый метод, основанный на комбинации нескольких деревьев решений. Модель строит множество деревьев решений, и итоговый результат вычисляется путём усреднения результатов всех деревьев (для регрессии) или голосования (для классификации).

Пример применения: Предсказание цены акций, классификация изображений.

Метод опорных векторов (Support Vector Machines, SVM):

Описание: Алгоритм, который ищет гиперплоскость, разделяющую данные на разные классы с максимальным зазором между ними. SVM эффективен для задач с малым количеством признаков.

Пример применения: Распознавание изображений, классификация текстов.

k-ближайших соседей (k-Nearest Neighbors, k-NN):

Описание: Простой алгоритм, который классифицирует новые данные, сравнивая их с k ближайшими примерами в обучающем наборе. Каждый новый объект получает класс, на который указывают большинство соседей.

Пример применения: Рекомендательные системы, прогнозирование покупок.

В обучении без учителя (Unsupervised Learning) данные не содержат явных меток или правильных ответов. Модель обучается находить скрытые структуры и закономерности в данных.

Кластеризация k-средних (k-Means Clustering):

Описание: Один из наиболее популярных алгоритмов кластеризации, который делит данные на k кластеров, минимизируя расстояния между точками в пределах кластера.

Пример применения: Сегментация клиентов, обнаружение аномалий.

Иерархическая кластеризация (Hierarchical Clustering):

Описание: Алгоритм, который строит дерево кластеров, объединяя или разделяя кластеры на основе расстояний между ними. Результатом является дендрограмма, которая показывает связи между объектами.

Пример применения: Биоинформатика, анализ данных пользователей.

Метод главных компонент (Principal Component Analysis, PCA):

Описание: Техника уменьшения размерности, которая преобразует исходные данные в набор новых переменных, которые линейно некоррелированы и охватывают максимальную дисперсию данных.

Пример применения: Уменьшение размерности данных для визуализации, сжатие изображений.

Ассоциативные правила (Association Rule Learning):

Описание: Алгоритмы, которые ищут закономерности в данных, такие как связь между покупками товаров. Основной алгоритм этого подхода — Apriori.

Пример применения: Анализ корзины покупок, выявление связанных событий.

Обучение с подкреплением (Reinforcement Learning) основывается на взаимодействии агента с окружающей средой. Агент выполняет действия,

получает за них вознаграждения или наказания и стремится максимизировать суммарное вознаграждение.

Q-Learning:

Описание: Модель использует функцию ценности для оценки действий и выбора тех, которые приведут к максимальной награде в долгосрочной перспективе.

Пример применения: Управление роботами, игра в компьютерные игры.

Deep Q-Networks (DQN):

Описание: Комбинация Q-Learning и глубокого обучения, где нейронные сети используются для предсказания наград в будущем.

Пример применения: Автономные транспортные средства, игры с элементами стратегии.

1.3 Алгоритмы глубокого обучения

Глубокое обучение — это подмножество машинного обучения, которое использует многослойные нейронные сети. В отличие от традиционных методов машинного обучения, глубокие нейронные сети способны извлекать сложные закономерности из больших объемов данных.

Основные архитектуры нейронных сетей

Полносвязные нейронные сети (Fully Connected Networks, FCN):

Описание: Базовая архитектура нейронной сети, где каждый нейрон связан с каждым нейроном следующего слоя. Эти сети используются для задач, связанных с табличными данными.

Пример применения: Прогнозирование финансовых показателей, классификация текстов.

Свёрточные нейронные сети (Convolutional Neural Networks, CNN):

Описание: CNN используются для обработки данных с пространственной структурой, таких как изображения. Эти сети используют свёрточные слои, которые автоматически извлекают признаки на разных уровнях абстракции.

Пример применения: Распознавание изображений, обработка видео.

Рекуррентные нейронные сети (Recurrent Neural Networks, RNN):

Описание: RNN используются для обработки последовательных данных, таких как тексты или временные ряды. Эти сети запоминают информацию о предыдущих шагах, что делает их эффективными для задач с временной зависимостью.

Пример применения: Машинный перевод, распознавание речи.

LSTM (Long Short-Term Memory):

Описание: Улучшенная версия RNN, которая решает проблему "затухающих градиентов", сохраняя важную информацию в долгосрочной перспективе.

Пример применения: Анализ временных рядов, прогнозирование цен.

Трансформеры (Transformers):

Описание: Современная архитектура, основанная на механизме внимания. Трансформеры способны эффективно обрабатывать длинные последовательности данных, как в задачах обработки текста, так и в компьютерном зрении.

Пример применения: GPT-3, BERT, машинный перевод.

Градиентный спуск (Gradient Descent):

Описание: Основной алгоритм оптимизации, который минимизирует функцию потерь, обновляя веса сети в направлении антиградиента. Используется практически во всех нейронных сетях.

Пример применения: Настройка параметров нейронных сетей.

Адаптивный градиент (Adam):

Описание: Популярный алгоритм оптимизации, который сочетает в себе моменты и адаптивную скорость обучения, что позволяет ускорить обучение и улучшить сходимость.

Пример применения: Большие нейронные сети, такие как CNN и RNN.

1.4 Автоматическое обнаружение объектов

Автоматическое обнаружение объектов — это одна из ключевых задач в области компьютерного зрения, которая включает в себя не только распознавание объектов на изображениях или видео, но и их локализацию в пространстве кадра. Благодаря стремительному развитию методов машинного и глубокого обучения, автоматическое обнаружение объектов стало одной из самых востребованных технологий в различных отраслях, включая безопасность, робототехнику, автономные транспортные средства и медицинские приложения.

Автоматическое обнаружение объектов — это процесс идентификации объектов различных категорий (например, автомобили, люди, животные и т. д.) на изображении или видео и их точного локализования с помощью рамок, называемых ограничивающими рамками (bounding boxes). Данный процесс сочетает задачи классификации (определение, к какому классу принадлежит объект) и сегментации (выделение областей, где находятся объекты).

Существует несколько категорий алгоритмов, используемых для обнаружения объектов. Большинство современных методов базируются на глубоких нейронных сетях, таких как свёрточные нейронные сети (CNN), которые способны обрабатывать пространственные структуры данных.

До широкого распространения глубокого обучения обнаружение объектов осуществлялось с помощью более простых, но эффективных методов:

Алгоритм Viola-Jones:

Описание: Один из первых широко применяемых методов для обнаружения лиц, который использует каскадный классификатор на основе

признаков Хаара. Алгоритм хорошо подходит для задач реального времени и требует минимальных вычислительных ресурсов.

Пример применения: Обнаружение лиц в камерах видеонаблюдения.

Histogram of Oriented Gradients (HOG):

Описание: Этот алгоритм использует гистограммы направленных градиентов для описания формы и структуры объектов. Он особенно эффективен для распознавания пешеходов.

Пример применения: Обнаружение пешеходов в автомобильных системах.

Deformable Part Models (DPM):

Описание: DPM разделяет объект на несколько частей и обучает модель для их поиска в различных положениях и масштабах. Это позволяет алгоритму обнаруживать объекты, имеющие изменяемые формы (например, людей в разных позах).

Пример применения: Обнаружение сложных объектов, таких как транспортные средства или люди.

Современные методы автоматического обнаружения объектов основаны на глубоких свёрточных нейронных сетях (CNN) [3,10]. Эти модели обучаются на огромных наборах данных, что позволяет им автоматически извлекать сложные признаки и адаптироваться к различным типам объектов и условий.

R-CNN (Regions with Convolutional Neural Networks):

Описание: Один из первых методов, предложенных для обнаружения объектов с использованием CNN. R-CNN выполняет сегментацию изображения на регионы интереса (ROI), а затем классифицирует и локализует объекты в этих регионах с помощью CNN.

Пример применения: Обнаружение транспортных средств, людей или животных на изображениях.

Fast R-CNN и Faster R-CNN:

Описание: Улучшенные версии оригинального R-CNN, которые значительно ускоряют процесс обнаружения. Fast R-CNN интегрирует этапы классификации и локализации, а Faster R-CNN добавляет специальную сеть для быстрого генерации регионов интереса.

Пример применения: Системы безопасности и мониторинга, автоматическое вождение.

YOLO (You Only Look Once):

Описание: Один из самых быстрых и эффективных методов для обнаружения объектов. YOLO рассматривает изображение целиком и делит его на сетку, в каждой ячейке которой пытается предсказать наличие объекта и его координаты [9]. Это позволяет YOLO работать в реальном времени.

Пример применения: Автономные автомобили, дроны, системы видеонаблюдения.

SSD (Single Shot MultiBox Detector):

Описание: Похож на YOLO, SSD также использует CNN для обнаружения объектов в различных масштабах и локализации их в одном проходе. SSD имеет преимущества в точности и скорости по сравнению с R-CNN.

Пример применения: Автоматическое отслеживание объектов в видео.

RetinaNet:

Описание: RetinaNet использует технику Focal Loss для улучшения точности обнаружения мелких объектов и объектов, которые трудно классифицировать. Этот алгоритм эффективен для сложных задач, где классы объектов могут быть несбалансированы.

Пример применения: Обнаружение объектов в медицинских изображениях, а также сложные сценки с малочисленными объектами.

Помимо обнаружения и классификации объектов, сегментация объектов позволяет точно выделить их границы на изображении. Существует два основных подхода:

Сегментация на уровне пикселей (Semantic Segmentation):

Описание: Сегментация объектов на уровне пикселей разделяет каждый пиксель изображения на категории (например, автомобиль, дорога, человек). В отличие от обнаружения объектов с использованием ограничивающих рамок, сегментация предоставляет более точные границы объектов.

Пример применения: Автономное вождение, медицинская диагностика.

Instance Segmentation:

Описание: Этот подход расширяет семантическую сегментацию, позволяя различать разные экземпляры одного и того же объекта (например, два разных автомобиля на одном изображении).

Пример применения: Обработка изображений для автономных роботов, 3D-сканирование.

1.5 Применение автоматического обнаружения объектов

Автоматическое обнаружение объектов на основе ИИ находит применение во множестве отраслей, предоставляя высокую точность и скорость обработки данных:

Автономные транспортные средства:

Обнаружение и классификация пешеходов, транспортных средств, дорожных знаков и препятствий на пути автомобиля. Системы, такие как Tesla Autopilot, используют глубокие нейронные сети для анализа данных с камер и сенсоров, чтобы принимать решения в режиме реального времени.

Безопасность и видеонаблюдение:

Камеры видеонаблюдения с ИИ могут автоматически идентифицировать подозрительные объекты или людей, отслеживать их передвижения и уведомлять оператора о возможных угрозах. Это используется для предотвращения преступлений, контроля доступа и мониторинга общественной безопасности.

Медицина:

В медицинских изображениях (например, рентген, МРТ, КТ) алгоритмы обнаружения объектов могут находить аномалии, такие как опухоли или патологические изменения. Это повышает точность диагностики и помогает врачам в раннем обнаружении заболеваний.

Робототехника:

Роботы с системами обнаружения объектов могут взаимодействовать с окружающей средой, идентифицировать и манипулировать объектами. Это критически важно для промышленных роботов, дронов и домашних роботов-помощников.

Обработка изображений в биометрии:

Алгоритмы распознавания лиц используются для аутентификации пользователей в системах безопасности, разблокировки устройств и анализа данных видеонаблюдения.

Сельское хозяйство:

Системы обнаружения объектов могут анализировать состояние растений, выявлять болезни или сорняки, а также автоматически управлять сельскохозяйственной техникой для оптимизации обработки полей.

1.6 Принципы работы систем обнаружения

Системы автоматического обнаружения объектов построены на основе нескольких ключевых принципов, которые позволяют моделям извлекать важную информацию из данных и принимать решения на основе этой информации. Эти принципы обеспечивают высокую точность распознавания объектов, особенно в таких сложных условиях, как изменение освещения, сложные фоны и наличие шумов на изображениях.

Существуют несколько методов, используемых для автоматического обнаружения и классификации объектов. Среди наиболее популярных методов — метод опорных векторов (SVM) и конволюционные нейронные сети (CNN).

Эти методы можно применять как для обнаружения, так и для классификации объектов.

Метод опорных векторов (SVM) — это один из наиболее популярных алгоритмов машинного обучения, используемых для задач классификации и регрессии. Основной идеей метода является нахождение гиперплоскости, которая максимально разделяет два (или более) класса объектов.

Принцип работы: Алгоритм строит гиперплоскость в многомерном пространстве таким образом, чтобы разделить объекты, относящиеся к разным классам, с максимальным зазором. Объекты, которые находятся на границах разделяющей плоскости, называются опорными векторами. Эти векторы определяют окончательное положение гиперплоскости.

Пример применения: SVM активно используется в биометрии для распознавания лиц и отпечатков пальцев, в анализе текстов для классификации и фильтрации спама, а также в распознавании изображений.

Конволюционные нейронные сети (CNN) являются основным методом для обработки данных с пространственной структурой, таких как изображения. В отличие от традиционных нейронных сетей, CNN способны автоматически извлекать признаки с различных уровней абстракции, что делает их особенно эффективными для задач обнаружения и классификации объектов.

Принцип работы: CNN используют свёрточные слои, которые фильтруют изображение, чтобы извлечь важные признаки. Затем эти признаки передаются на полносвязные слои для окончательной классификации. Каждый свёрточный слой обучается извлекать более сложные признаки (например, грани, текстуры, формы).

Пример применения: CNN активно используются для распознавания объектов в реальном времени (например, пешеходов, транспортных средств), в медицинских приложениях для анализа рентгеновских снимков и МРТ, а также в системах видеонаблюдения и робототехнике.