

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»

Кафедра информатики и программирования

Реферат

КОМПЬЮТЕРНОЕ ЗРЕНИЕ

студента 1 курса 151 группы
направления 09.03.04 — Программная инженерия
факультета компьютерных наук и технологий
Савельева Даниила Витальевича

Саратов 2023

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	3
1 КОМПЬЮТЕРНОЕ ЗРЕНИЕ КАК ОБЛАСТЬ НАУКИ	4
1.1 ЧТО ТАКОЕ КОМПЬЮТЕРНОЕ ЗРЕНИЕ?	4
1.2 ИСТОРИЯ КОМПЬЮТЕРНОГО ЗРЕНИЯ	4
2 КАК РАБОТАЕТ КОМПЬЮТЕРНОЕ ЗРЕНИЕ	6
2.1 ЗАДАЧА КОМПЬЮТЕРНОГО ЗРЕНИЯ И ЕЁ ДОСТИЖЕНИЕ	6
2.2 РАСПОЗНОВАНИЕ ИНФОРМАЦИИ	6
3 ЭФФЕКТИВНОСТЬ КОМПЬЮТЕРНОГО ЗРЕНИЯ	8
3.1 ДОСТОИНСТВА И НЕДОСТАТКИ КОМПЬЮТЕРНОГО ЗРЕНИЯ	8
3.2 ПРОБЛЕМЫ КОМПЬЮТЕРНОГО ЗРЕНИЯ	8
4 КОМПЬЮТЕРНОЕ ЗРЕНИЕ В ЖИЗНИ	10
4.1 ДЛЯ ЧЕГО ИСПОЛЬЗУЕТСЯ КОМПЬЮТЕРНОЕ ЗРЕНИЕ	10
4.2 ТЕНДЕНЦИИ КОМПЬЮТЕРНОГО ЗРЕНИЯ	11
4.3 ПРИМЕРЫ ИСПОЛЬЗОВАНИЯ КОМПЬЮТЕРНОГО ЗРЕНИЯ ...	12
ЗАКЛЮЧЕНИЕ	14
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	15

ВВЕДЕНИЕ

В современном мире, вслед за возрастающими потребностями человека, появляются новые технологии и развиваются старые. Одной из самых перспективных и быстро развивающихся информационных технологий является компьютерное зрение.

Целью моей работы является изучение того, почему компьютерное зрение является инновационным направлением, потенциально имеющим большое будущее.

Для достижения цели мне предстоит выполнить ряд задач:

- узнать, что такое компьютерное зрение;
- определить какие основные задачи стоят перед этой областью науки и как она их достигает;
- выделить достоинства и недостатки компьютерного зрения
- разузнать какие возможности предоставляет компьютерное зрение для автоматизации процессов и оптимизации работы различных сфер деятельности.

1 КОМПЬЮТЕРНОЕ ЗРЕНИЕ КАК ОБЛАСТЬ НАУКИ

1.1 ЧТО ТАКОЕ КОМПЬЮТЕРНОЕ ЗРЕНИЕ?

Зрение является основным источником информации для человека о мире, но его сложность стала явной только после попыток воспроизвести его на компьютере. Люди приступили к разработкам искусственного зрения, то есть переносу возможности «видеть», как человек, на компьютер, и на этом пути возникли огромные трудности, связанные с различиями в естественных изображениях одного и того же объекта и отсутствием трехмерной информации на двумерных изображениях. Решение этих проблем привело к возникновению области наук и технологий, которая получила название компьютерного зрения.

Компьютерное зрение — это молодая, но быстро развивающаяся отрасль, которая работает над разработкой систем, которые могут анализировать информацию на изображениях и видео для использования в приложениях в различных областях жизни, используя для этого вычислительно-математические методы и технологии обработки изображений.

1.2 ИСТОРИЯ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Когда именно возникла идея компьютерного зрения, точно неизвестно. Считается, что одной из первых статей в научном журнале, так или иначе затрагивающих эту тему, была публикация «Receptive fields of single neurons in the cat's striate cortex». Её написали в 1959 году Дэвид Хьюбел и Торстен Визель, нейропсихологи из медицинского института Уилмера в США. В своей работе они изучали свойства нейронов зрительной коры кошек. И заметили, что зрительный опыт способен влиять на эти нейроны.

В это же время в схожем направлении думал и профессор Массачусетского технологического института Лоуренс Робертс. Он был уверен, что можно научить компьютер видеть, создал систему распознавания форм предметов и защитил в МТИ докторскую диссертацию по этому направлению.

И действительно, очень скоро компьютеры научились распознавать статичные изображения, в 80-х годах появились теории систем распознавания движущихся объектов на видео, а в 90-х учёные работали над прототипами беспилотных автомобилей [1].

Люди поняли, что технологии компьютерного зрения могут использоваться везде, где есть какие-либо изображения. А благодаря развитию интернета

появились массивы оцифрованных изображений, которые можно было анализировать. Это позволило обучать компьютеры. Сначала они могли лишь распознавать текст на сканах документов, но постепенно задачи усложнялись. Так технологии дошли до обнаружения лиц и их почти мгновенного распознавания на фото и видео.

Из истории компьютерного зрения можно сделать вывод, что эта область науки сочетает в себе знания из множества дисциплин, таких как математика, физика, биология, информатика и других.

2 КАК РАБОТАЕТ КОМПЬЮТЕРНОЕ ЗРЕНИЕ

2.1 ЗАДАЧА КОМПЬЮТЕРНОГО ЗРЕНИЯ И ЕЁ ДОСТИЖЕНИЕ

Основная задача компьютерного зрения— это перевод изображения в понятную для компьютера форму. Эта задача включает в себя различные этапы обработки изображения, такие как выделение контуров, сглаживание изображения, разделение на отдельные объекты, определение их размера, формы и положения. Обработка изображения может быть очень сложной задачей, так как на изображениях может быть много различных объектов, которые должны быть обработаны.

Специалисты ищут всё более совершенные способы обучить компьютер правильно видеть и извлекать информацию из увиденного. Казалось бы, это так просто: научить искусственный интеллект распознавать визуальные образы. Но не тут-то было.

Компьютер видит не так, как люди. У него нет нашего жизненного опыта и способности так же легко идентифицировать объекты на изображении. Он не способен отличить дом от дерева, не имея каких-то исходных данных. Чтобы научить компьютер видеть и понимать, что находится на изображении, люди используют технологии машинного обучения.

Для этого собирают большие базы данных, из которых формируют дата-сеты. Выделив признаки и их комбинации для идентификации похожих объектов, можно натренировать модель машинного обучения распознавать нужные типы закономерностей. Конечно, даже после загрузки нескольких дата-сетов модели могут неверно распознавать некоторые объекты. Если такое случается, модели дообучают на новых наборах данных [1].

2.2 РАСПОЗНОВАНИЕ ИНФОРМАЦИИ

Любая система компьютерного зрения при анализе изображения проходит через три основных этапа:

- Классификация изображения.

Компьютер присваивает картинке некий класс из заранее известных.

- Локализация конкретного объекта.

Чем больше объектов на изображении, тем сложнее требуется нейросеть.

- Построение изображения.

Программа убирает шум и повышает качество картинки, выделяя доми-

нирующие и малозначительные объекты [1].

Каждый этап сложен, ведь компьютер не понимает, что на изображении важное, а что не очень. Поэтому сначала картинка проходит через внутренние алгоритмы, заложенные разработчиками. Так компьютер получает начальные сведения об изображении. Затем компьютер находит объекты и их границы. Для этого есть разные способы. Например, можно использовать размытие по Гауссу, когда изображение размывают несколько раз, выявляя самые контрастные фрагменты. Эти значимые места в дальнейшем считаются объектами.

Значимые места компьютер при помощи любого из популярных алгоритмов (например, SIFT, SURF, HOG) описывает в числовом виде. Такая запись называется дескриптором. Она позволяет полно и точно сравнить фрагменты изображений, не используя сами фрагменты. Но так как сравнение — это тяжёлая вычислительная операция, то дескрипторы кластеризуют: делят на группы. В каждой группе находятся дескрипторы разных изображений, но с общими характеристиками.

Кластеризация и следующее за ним квантование (обобщение) дескрипторов уменьшает объём данных, которые приходится обрабатывать компьютеру. А заодно позволяет быстрее распознавать объекты и сравнивать изображения [1].

3 ЭФФЕКТИВНОСТЬ КОМПЬЮТЕРНОГО ЗРЕНИЯ

3.1 ДОСТОИНСТВА И НЕДОСТАТКИ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Компьютерное зрение обладает рядом очень важных преимуществ:

- Высокая точность и скорость работы.
- Возможность обработки большого объема данных.
- Автоматизация процессов обработки изображений и видео.
- Снижение затрат на труд и ресурсы.
- Минимизация ошибок, связанных с человеческим фактором.
- Решение сложных задач в медицине, автоматизации производства, робототехнике и других областях.

Однако, несмотря на все преимущества, компьютерное зрение имеет свои ограничения. Некоторые области, такие как неопределенность в яркостных данных, могут быть сложны для алгоритмов компьютерного зрения, особенно в условиях низкого освещения. Также алгоритмы компьютерного зрения могут быть чувствительны к шуму на изображении.

3.2 ПРОБЛЕМЫ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Современные разработки в области компьютерного зрения сталкиваются с рядом сложностей, одной из которых является недостаточное количество исходных данных. Несмотря на распространенность фото и видеоаппаратуры, у дата-сайентистов не всегда есть достаточное количество материалов для обучения алгоритмов, это может быть связано с законодательными и этическими ограничениями, а также географическими препятствиями. Низкое качество материалов для обучения является еще одной проблемой, так как это может привести к ошибкам в процессе обработки данных. Большая часть материалов для обучения требует ручной разметки данных, что является сложным, долгим, монотонным процессом, в котором человеческий фактор играет важную роль, что может привести к ошибкам в датасетах.

В апреле 2021 года ученые из Массачусетского технологического института выяснили, что 5,8% изображений одного из самых популярных тестовых наборов данных ImageNet подписаны неправильно. Среди самых распространенных ошибок — неправильные подписи объектов: на фотографиях гриб может быть отмечен как ложка, а лягушка — кошкой [2].

Тестовые датасеты могут содержать ошибки, которые влияют на качество работы алгоритмов машинного обучения, поэтому разработчикам алгоритмов ИИ необходимо более тщательно обрабатывать данные при создании моделей. Чтобы обработать большие объемы медиаданных, требуются дорогостоящие вычислительные ресурсы, и хотя облачные сервисы могут частично решить эту проблему, все еще необходимо стабильное широкополосное интернет-соединение, особенно для обработки реального времени. Решить эту проблему могут граничные вычисления или edge computing, которые осуществляются непосредственно в местах сбора данных, используя одноплатные компьютеры или вычислительные процессоры, оборудованные ИИ-алгоритмами. Однако, несмотря на недорогие цены на одноплатные компьютеры, они все еще не обладают достаточной мощностью для обработки больших объемов данных, особенно в режиме реального времени.

4 КОМПЬЮТЕРНОЕ ЗРЕНИЕ В ЖИЗНИ

4.1 ДЛЯ ЧЕГО ИСПОЛЬЗУЕТСЯ КОМПЬЮТЕРНОЕ ЗРЕНИЕ

Компьютерное зрение представляет собой мощную технологию, которую можно использовать в комбинации с различными сенсорными устройствами и приложениями для применения в множестве практических ситуаций. Оно может использоваться для:

- Упорядочения содержимого

С помощью компьютерного зрения можно автоматически идентифицировать и классифицировать людей или объекты на фотографиях, используя различные данные. Подобные приложения для распознавания широко применяются в системах хранения фотографий и социальных сетях.

- Извлечения текста

Оптическое распознавание символов можно использовать, чтобы упростить обнаружение содержимого в данных с большим объемом текста и реализовать обработку документов в сценариях автоматизированной обработки.

- Дополненной реальности

Компьютерное зрение обнаруживает и отслеживает физические объекты в реальном времени. Эти сведения затем используются для реалистичного размещения виртуальных объектов в физической среде.

- Сельского хозяйства

Изображения посевов, сделанные со спутников, дронов или самолетов, можно анализировать для сбора данных об урожае, обнаружения зарослей сорняков или определения дефицита питательных веществ.

- Автономных транспортных средств

Беспилотные автомобили используют идентификацию и отслеживание объектов в реальном времени для сбора данных о ситуации вокруг автомобиля и построения маршрута.

- Здравоохранения

Фотографии или изображения, созданные медицинскими устройствами, можно анализировать, чтобы упростить и ускорить врачам выявление проблем и постановку диагнозов и сделать их более точными.

- Спорта

Обнаружение и отслеживание объектов используются для анализа игры и

корректировки стратегии.

- **Производства**

Компьютерное зрение может отслеживать производственное оборудование с целью обслуживания. Его также можно применять для контроля качества продукции и упаковки на производственных линиях.

- **Пространственного анализа**

Система идентифицирует людей или объекты, такие как автомобили, в пространстве и отслеживает их передвижение.

- **Распознавания лиц**

Компьютерное зрение может применяться для идентификации людей [3].

- **Изучения космоса**

Снимки со спутников и телескопов долгое время анализировали люди. За многолетнюю историю наблюдения за космосом накопилось огромное количество данных из самых разных источников. И в этих данных может содержаться немало ценной информации, которую просто не заметили. Например, в декабре 2017 года астрономы, используя искусственный интеллект, проанализировали данные, собранные телескопом «Кеплер». Компьютерное зрение увидело то, что не заметил человеческий глаз: солнечную систему с восемью планетами [1].

4.2 ТЕНДЕНЦИИ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Одним из главных направлений в области компьютерного зрения являются генеративно-сопоставительные нейросети (GAN). В последнее время эти алгоритмы используются не просто для стилизации фотографий и видео под картины известных художников, но и для создания качественных подделок.

Например, проект *This Person Does not Exist* использует GAN для генерирования фотореалистичных изображений людей, которых на самом деле не существует. По схожему принципу работают и другие проекты: алгоритм по созданию ненастоящих котов *This Cat Does not Exist*, или кроссовок — *This Sneaker Does not Exist* [2].

Подобные алгоритмы позволяют исследователям и разработчикам создавать синтетические наборы данных для обучения моделей. Такие датасеты легче собрать и они решают некоторые правовые и этические вопросы использования изображений.

Другим важным направлением в области является моделирование 3D-

сцен. Для реализации данной задумки разрабатываются специальные алгоритмы, которые, используя серию фотографий с разных ракурсов, способны воссоздать сцену в трехмерном пространстве.

Эту технологию активно используют в строительстве, робототехнике, анимации, дизайне интерьеров и военном деле.

Исследователи отмечают, что на сегодня алгоритмам тяжело воспроизводить сложные текстуры, например, листьев на деревьях. Тем не менее в ближайшем будущем такие инструменты смогут значительно упростить работу 3D-дизайнерам [2].

4.3 ПРИМЕРЫ ИСПОЛЬЗОВАНИЯ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Сейчас многие компании занимаются развитием и совершенствованием компьютерного зрения. Создаются полезные приложения, сервисы, сайты, проектируются инновационные устройства. Так примеры использования этой технологии становятся все более и более интересными, а иногда и очень необычными.

Например, в области автомобилей и беспилотного пилотирования компания NVIDIA создала суперкомпьютер NVIDIA Drive 2, который уже служит базовой платформой для беспилотников Tesla, Volvo, Audi, BMW и Mercedes-Benz, а также технологию искусственного восприятия DriveNet представляющую собой самообучаемое компьютерное зрение, работающее на основе нейронных сетей. С ее помощью лидары, радары, камеры и ультразвуковые датчики способны распознавать окружение, дорожную разметку, транспорт и многое другое. Благодаря этим технологиям крупные автопроизводители смогут представить полностью автономные машины.

Компьютерное зрение способно усовершенствовать систему умный дом, сделав нашу жизнь более комфортной и удобной:

- Устройство eyeSight's Singlecue Gen 2 использует компьютерное зрение (распознавание жестов, анализ лица, определение действий) и позволяет управлять с помощью жестов телевизором, «умной» системой освещения и холодильниками.
- Краудфандинговый проект Науо, пожалуй, является самым интересным новым интерфейсом. Эта технология позволяет создавать виртуальные средства управления по всему дому — просто подняв или опустив руку, вы можете увеличить или уменьшить громкость музыки, или же включить

свет на кухне, взмахнув рукой над столешницей. Все это работает благодаря цилиндрическому устройству, использующему компьютерное зрение, а также встроенную камеру и датчики 3D, инфракрасного излучения и движения.

- Устройство FridgeCam от компании Smarter крепится к стенке холодильника и может определять, когда истекает срок годности, сообщать, что именно находится в холодильнике, и даже рекомендовать рецепты блюд из выбранных блюд.

Технология компьютерного зрения совершенствует робототехнику:

- Alexa от Amazon, Google Home и прочие цифровые помощники и роботы, доступные на рынке, вроде LG Hub и Kuri от Mayfield Robotics, обладают базовыми навыками компьютерного зрения и могут определить, кто с ними разговаривает, или же выгнать собаку с дивана.
- Компания ITRI разработала систему Intelligent Vision System, которая использует глубинное обучение и компьютерное зрение, чтобы роботы могли различать объекты разного размера (фигурки, чашки) и определять их положение. Распознав объект, робот сможет взять его и принести в нужное место. Такие навыки отлично быгодились для обслуживания столиков в ресторане или для игры в шахматы [4].
- Компания Google работает над технологией, которая позволяет определять калорийность различных блюд по их фотографиям. Эта технология называется Im2Calories, и в ней используются предварительные наработки DeepMind, стартапа, который Google приобрела в прошлом году.

Im2Calories определяет размер порции, вычисляя массу на основании того, какой объем тарелки занимает блюдо. Технологии все равно, с каким разрешением пользователь делает фото — высоким или низким, результат в обоих случаях получается примерно одинаковый.

И это только малая часть примеров использования этой удивительной технологии.

ЗАКЛЮЧЕНИЕ

Компьютерное зрение – это высокотехнологичная область науки, предлагающая широкий спектр приложений во многих областях, от медицины и робототехники до производства и охраны правопорядка. Развитие этой области науки и технологии открывает новые возможности для повышения эффективности работы различных организаций и улучшения качества жизни людей, а в будущем оно будет иметь еще больший потенциал. Именно поэтому так важно сейчас заниматься разработкой и совершенствованием компьютерного зрения. В своей работе я доказал актуальность и востребованность этой области науки.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 Компьютерное зрение: от распознавания текста до изучения космоса [Электронный ресурс].— URL: <https://cloud.yandex.ru/blog/posts/2022/05/computer-vision> (дата обращения: 30.04.2023). Загл. с экр. Яз. рус.
- 2 Что такое компьютерное зрение? (машинное обучение) [Электронный ресурс].— URL: <https://forklog.com/cryptorium/ai/chto-takoe-kompyuternoe-zrenie-mashinnoe-obuchenie> (дата обращения: 01.05.2023). Загл. с экр. Яз. рус.
- 3 Что собой представляет компьютерное зрение? [Электронный ресурс].— URL: <https://azure.microsoft.com/ru-ru/resources/cloud-computing-dictionary/what-is-computer-vision/> (дата обращения: 03.05.2023). Загл. с экр. Яз. рус.
- 4 8 примеров использования компьютерного зрения [Электронный ресурс].— URL: <https://rb.ru/list/computer-vision-in-practice/> (дата обращения: 16.05.2023). Загл. с экр. Яз. рус.
- 5 Потапов А.С. Системы компьютерного зрения. Учебное пособие. – СПб: Университет ИТМО, 2016. – 161 с.
- 6 Шапиро Л., Стокман Д.Ж. Компьютерное зрение. - 1-е изд. - М.: Лаборатория знаний, 2020. - 763 с.
- 7 Компьютерное зрение: технологии, рынок, перспективы [Электронный ресурс].— URL: <https://www.tadviser.ru/index.php/Статья:Компьютерноезрение:технологии,рынок,п> (дата обращения: 27.04.23). Загл. с экр. Яз. рус.
- 8 Как устроено компьютерное зрение [Электронный ресурс].— URL: <https://practicum.yandex.ru/blog/chto-takoe-kompyuternoe-zrenie/> (дата обращения: 28.04.23). Загл. с экр. Яз. рус.
- 9 Селянкин В.В. Компьютерное зрение. Анализ и обработка изображений: учебное пособие для вузов / В. В. Селянкин. — 2е изд., стер. — Санкт-Петербург : Лань, 2021. — 152 с.

- 10 Компьютерное зрение: направления и сферы применения [Электронный ресурс].— URL: <https://www.jetinfo.ru/computer-vision-technology-review/> (дата обращения: 09.05.23). Загл. с экр. Яз. рус.
- 11 Смотри внимательно Как компьютеры видят мир и зачем это нужно [Электронный ресурс].— URL: <https://nplus1.ru/material/2022/10/10/computer-vision> (дата обращения: 04.05.23). Загл. с экр. Яз. рус.