

ОБУЧЕНИЕ СВЁРТОЧНОЙ НЕЙРОННОЙ СЕТИ НА ТРЁХМЕРНЫХ ОБЪЕКТАХ

Научная статья

Лащик Н.М.^{1,*}

¹МИРЭА - Российский технологический университет, Москва, Российская Федерация

* Корреспондирующий автор (lkolya97[at]gmail.com)

Аннотация

В данном исследовании изучается подход к обучению сверточных нейронных сетей с использованием трехмерной компьютерной графики для представления набора данных. Сверточные нейронные сети обучаются и классифицируют объекты, используя ортогональные проекции трехмерных объектов. Одной из сложностей, возникающих в процессе классификации, является зеркальное отображение объекта, которое можно преодолеть с помощью коэффициента корреляции Пирсона. В ходе работы был проведен анализ архитектуры существующих решений для эффективного распознавания изображений и записи визуальной информации, которые могут быть использованы в качестве основы для сверточной нейронной сети. Для обеспечения корректной работы подготовленной нейронной сети были созданы объекты трехмерной графики, и после получения необходимой трехмерной модели, было создано черно-белое изображение. Именно на этих черно-белых изображениях обучается программа для распознавания объектов. Экспериментально доказано успешное функционирование нейронной сети. Таким образом, становится возможным распознавание реальных объектов на основе сверточных нейронных сетей, подготовленных в виртуальном окружении.

Ключевые слова: трёхмерный объект, распознавание категорий объектов, зеркальное отображение объектов, сверточные нейронные сети, машинное обучение.

TRAINING A CONVOLUTIONAL NEURAL NETWORK ON THREE-DIMENSIONAL OBJECTS

Research article

Lashchik N.M.^{1,*}

¹MIREA - Russian Technological University, Moscow, Russian Federation

* Corresponding author (lkolya97[at]gmail.com)

Abstract

This research examines an approach to training convolutional neural networks using three-dimensional computer graphics to represent a dataset. Converged neural networks are trained and classify objects using orthogonal projections of three-dimensional objects. One of the difficulties encountered in the classification process is object mirroring, which can be overcome by using Pearson correlation coefficient. This work analysed the architecture of existing solutions for efficient image recognition and visual information recording, which can be used as a basis for a convolutional neural network. To ensure the correct operation of the trained neural network, 3D graphics objects were created and after obtaining the required 3D model, black and white images were generated. It is on these black and white images that the object recognition software is trained. The successful functioning of the neural network has been experimentally proved. Thus, it becomes possible to recognize real objects based on convolutional neural networks trained in a virtual environment.

Keywords: three-dimensional object, object category recognition, object mirroring, convolutional neural networks, machine learning.

Введение

Идентификация объектов является одной из основных когнитивных способностей человеческого мозга. Сначала процесс идентификации начинается с восприятия объекта органами чувств, в основном зрением. Зрительное восприятие позволяет человеку собирать информацию о характерных особенностях наблюдаемых объектов [1].

В контексте искусственного интеллекта проблема идентификации объектов решается с помощью компьютерного зрения на основе сверточных нейронных сетей. Они позволяют манипулировать двухмерной графикой и трансформировать ее. Иногда, однако, может быть трудно своевременно предоставить обучающие данные. Люди, например, легко адаптируются к новым ситуациям и окружениям, наблюдая за новыми классами объектов. Для переноса автоматического устройства в новую среду часто требуется полное изменение новой базы знаний. Ясно, что многие существенные вопросы в сфере компьютерного зрения уже решены, но все еще есть задачи, требующие улучшения. В этой статье рассматриваются вопросы обучения сверточным нейронным сетям на базе трехмерной компьютерной графики.

Описание задачи

Идентификация объектов с использованием крайне ограниченных обучающих данных крайне важна для множества приложений компьютерного зрения. Эта проблема становится ещё более сложной, если системе нужно изучать новые категории объектов при наличии очень малого количества примеров. Использование сверточных нейросетей для решения таких задач кажется сложным, так как они требуют больших объёмов данных и страдают от «катастрофической забывчивости». Однако эта проблема может быть решена через подход обучения и идентификации на основе примеров: изучение категорий рассматривается как процесс изучения имеющихся образцов категории, то есть категория представлена лишь через известные экземпляры., $C \leftarrow \{O_1, \dots, O_n\}$, где C представление типа объекта, основанное на объединении слоев [2].

Обучение на основе примеров – также известное как память-ориентированное обучение – представляет собой базовый подход к оценке объектных представлений. Преимущество подхода, основанного на примерах, перед другими методами машинного обучения заключается в возможности быстрой адаптации объектной категориальной модели к ранее невиданному примеру путём сохранения нового экземпляра или удаления старого. Такое обучение позволяет идентифицировать объекты, используя лишь несколько экспонентных экземпляров, при этом сохраняя слишком много избыточных ценностей, что ведёт к значительному потреблению памяти и снижению скорости идентификации [3]. Поэтому каждый раз при появлении нового запроса проверяются его сохранённые ранее данные и целевому объекту присваивается ценность для нового экземпляра.

Работы по данной теме

В настоящее время для решения проблемы недостатка обучающих данных применяются следующие методы. Первый и самый распространённый метод, способный помочь в борьбе с недостатком обучающих данных – это полное отсутствие проверочных выборок. Этот метод наименее эффективен для решения данной проблемы, но имеет преимущества в виде простоты исполнения и экономии времени. К сожалению, последствия такого решения сложно предсказать, а в случае сложных или слабо понятных моделей от такого подхода лучше отказаться вовсе.

Второй метод заключается в использовании теории и методов нечёткой логики и нечётких множеств. Основная идея заключается в создании правил «если-то» для определения корреляций между различными входными переменными, что помогает решить проблему рассогласованности данных и их малого количества. С помощью этих правил и методов создаются новые примеры, помогающие увеличить доступный объём данных для обучения. Недостатком такого подхода может считаться то, что он требует значительных вычислительных ресурсов и опыта работы с нечёткими множествами.

Третий метод предлагает справляться с проблемой недостатка обучающих данных посредством разделения входных и целевых переменных на более маленькие подгруппы. Затем к каждому из этих сегментов применяются более простые искусственные нейросети – однослойные перцептроны. Такие сети могут быть обучены с использованием доступных объёмов данных. По окончании обучения перцептроны соединяются в одну структуру, известную как перцептронный ансамбль. Однако недостатком данного подхода может стать снижение точности модели, так как разделение переменных на подгруппы может вызвать потерю информации и уменьшение детализации модели. Более того, процесс разбиения переменных на сегменты и выбора оптимального числа подгрупп может быть достаточно сложным и требовать множества экспериментов и тестирований.

Предлагаемый подход

В последнее время большое внимание с точки зрения компьютерного зрения уделяется методу углублённого изучения, которое можно разделить на три категории в зависимости от входных данных: объёмные данные, в которых объект представляется как трёхмерная воксельная сетка, которая затем подаётся в сверточную нейронную сеть как входные данные об объекте; данные, базирующиеся на форме, в которых объект представлен в виде набора двумерных изображений, полученных путём проецирования объекта на плоскость; а также данные, основанные на массиве точек, в которых объект представлен предварительно размеченным трёхмерным изображением. Из этих методов, наиболее эффективными для распознавания объектов оказались методы, основанные на форме. Они позволили исследователям добиться наилучших результатов в распознавании на сегодняшний день. Вместе с тем, такое представление значительно облегчает изучение объекта в программе для создания трёхмерной компьютерной графики Blender.

Далее, множество точек виртуального объекта представлено набором точек p_i , где $i \in \{1, \dots, n\}$. Затем каждая точка может быть описана координатами x , y и z . Основываясь на анализе собственных векторов, создаются три основные оси объекта, а центр объекта вычисляется с помощью инструментов программного обеспечения Blender. Исходя из этого, могут быть созданы ортогональные проекции объекта для различных видов. Можно создать до шести типов таких проекций, однако для представления трёхмерного объекта достаточно трёх, включая виды спереди, сверху и сбоку. (Рис.1).

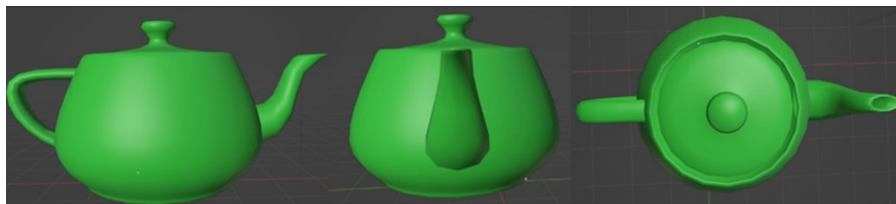


Рисунок 1 - Пример объекта и его проекций

Следует отметить, что направление собственных векторов не является однозначным, ортогональные проекции могут быть зеркально отражены. Для устранения неоднозначности используется коэффициент корреляции Пирсона. Тогда проецированная точка $\rho = (\alpha, \beta) \in \mathbb{R}^2$, где α – расстояние, перпендикулярное горизонтальной оси, а β – расстояние, перпендикулярное вертикальной оси. Коэффициент корреляции Пирсона r_x рассчитывается для направления оси X по формуле:

$$r_x = \frac{\sum(\alpha_i \beta_i - \bar{\alpha} \bar{\beta})}{\sqrt{\sum(\alpha_i - \bar{\alpha})^2} \sqrt{\sum(\beta_i - \bar{\beta})^2}} \quad (1)$$

где $\bar{\alpha}$ и $\bar{\beta}$ – средние значения α и β . Согласно формуле, определяются свойства коэффициента корреляции r : корреляция Пирсона r изменяется в диапазоне от -1 до +1; значение корреляции Пирсона r указывает на то, насколько близко точки расположены к прямой линии.

В частности, если корреляция Пирсона $r=+1$, то корреляция абсолютно положительна, а если корреляция Пирсона $r=-1$, то корреляция абсолютно отрицательна. Аналогично рассчитывается показатель r_y для оси Y , используя плоскость YoZ . Далее знак осей определяется как $s=r_x r_y$, где s может быть как положительным, так и отрицательным. В случае отрицательного значения s три проекции должны быть зеркально отражены, в то время как в случае положительного значения s они не должны отображаться.

Когда отличия нового объекта со всеми ранее известными категориями превышают пороговое значение, автоматическое устройство делает вывод о том, что этот объект не принадлежит к известным категориям, и, таким образом, инициализирует новую категорию, помеченную как «category_m+1», где m – количество известных в данный момент категорий. Более того, в случае измерения сходства разница между двумя объектами может быть рассчитана с использованием различных функций расстояния. Например, для оценки сходства двух объектов используется критерий хи-квадрат Пирсона.

При обработке каждого изображения необходимо реализовать распознавание контуров объектов (Рис.2).

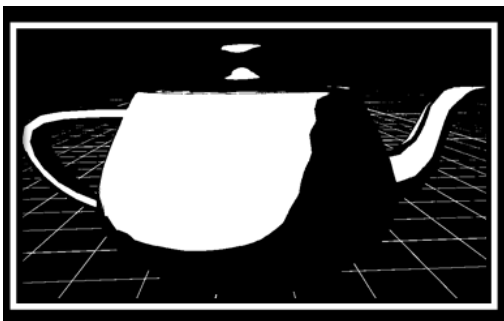


Рисунок 2 - Результат обработки изображения

В результате получается изображение (рис. 2). На нём все искомые объекты будут представлены белыми пятнами на полностью чёрном фоне. Затем можно использовать функцию `findContours` из библиотеки `OpenCV` для получения контуров этих белых пятен.

Сверточная нейронная сеть – это особый вид искусственной нейросети, оптимизированный для эффективного распознавания объектов на изображении. Она моделирует некоторые особенности зрительной коры мозга, где были обнаружены два вида клеток: простые клетки, реагирующие на прямые линии под различными углами, и сложные клетки, реакция которых зависит от активации определённого набора простых клеток. Есть множество реализаций сетей такого типа, предназначенных для классификации объектов на изображениях и распознавания паттернов.

Среди протестированных сетей наиболее эффективными выглядят модели из группы YOLOv3: YOLOv3-320 и YOLOv3-416. Их сильными сторонами являются высокая производительность (>30 кадров в секунду) и точность, которая в среднем не ниже, чем у других (>50%). Модели основаны на платформе Darknet.

На вход сети поступает изображение размером 448x448. Если оно не соответствует этому размеру, то оно растягивается или сжимается до такого размера. Сеть содержит 106 слоёв и использует только 4 вида слоёв и 2 операции над ними (рис.3).

Первый вид – это свёрточный слой. Он создаёт карту объектов по входным данным и выделяет на изображении основные закономерности. Второй вид – это маршрутизирующий слой, который нужен для определения двух слоёв, результаты которых должны быть объединены. Третий вид – это супервыборный слой. Он увеличивает размерность выхода благодаря карте объектов со свёрточного слоя. Это нужно для правильного объединения слоёв. Четвёртый вид – это обнаружение слоя. Он используется для нахождения рамок окон. Операция объединения соединяет карты объектов двух слоёв, а операция добавления суммирует карты элементов по элементам.

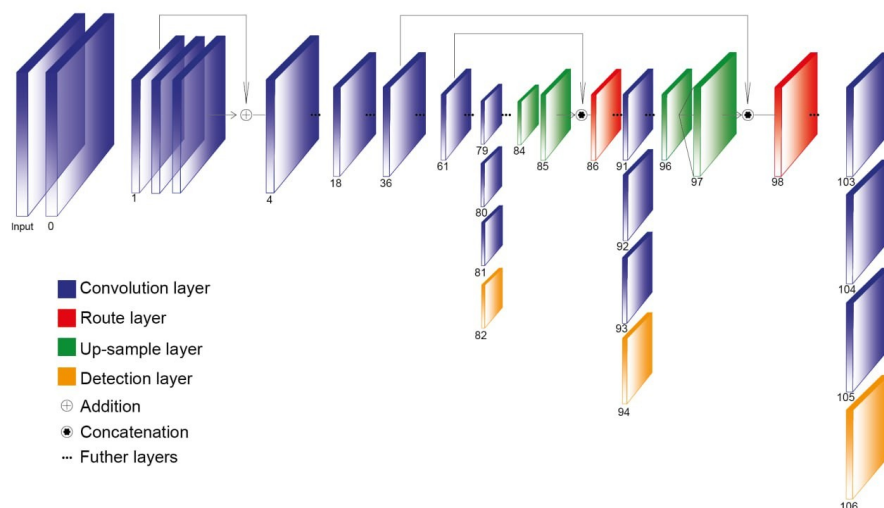


Рисунок 3 - Архитектура YOLOv3

Перед обучением необходимо настроить параметры сети. Для этого нужно установить количество классов равным исходному на уровнях обнаружения и на свёрточных уровнях, предшествующих им. Изменить размерность карты объектов на величину, равную количеству классов [4]. В каждую эпоху исходные изображения подвергаются различным модификациям: повороту изображения на малый угол, изменению цветового спектра и так далее. Для получения лучшего качества распознавания весов уже обученной нейросети использовались на наборе, созданном в программном обеспечении для создания трёхмерной графики «Blender». В качестве показателя точности YOLOv3 использует сумму среднеквадратических ошибок.

$$E_{MSE} = \frac{1}{n} \sum_{j=1}^n (d_j - y_j)^2 \quad (2)$$

В машинном обучении ошибка обучения – это разница между желаемым (целевым) значением d_j и фактическим значением y_j , которое получено на выходе модели при использовании обучающего набора данных с количеством наблюдений n .

При создании автоматической системы компьютерного зрения недостаточно просто найти объекты на изображении. Поскольку система получает непрерывный поток видео, необходимо также научить её отслеживать движение найденных объектов. Здесь на помощь приходят алгоритмы отслеживания объектов. Алгоритм отслеживания объекта – это метод, который определяет положение движущегося объекта (или группы объектов) во времени. Отследить движущийся объект может быть сложно из-за больших объёмов данных, которые содержит видео (визуальная информация). Эта сложность ещё больше усугубляется необходимостью использования методов распознавания объектов, что само по себе уже является сложной задачей [5].

Наиболее популярный метод, когда перемещение объекта предсказуемо и незначительно — это трекер MedianFlow. Он работает следующим образом [6]:

- алгоритм разбивает изображение отслеживаемого объекта на небольшие фрагменты;
- для каждого фрагмента определяется новое положение на следующем кадре путем вычисления оптического потока между двумя изображениями;
- чтобы оценить новое положение объекта, трекер вычисляет медиану векторов смещения всех фрагментов изображения;
- чтобы оценить изменение размера отслеживаемого объекта, трекер вычисляет медиану расстояний между всеми фрагментами изображения.

Трекер отслеживает объекты как вперёд, так и назад во времени. Он сравнивает эти две траектории и выбирает ту из них, которая имеет минимальную ошибку (разницу) между двумя кадрами. В отличие от других трекеров, которые продолжают работу и пытаются отслеживать объект, даже когда тот «потерян» или вышел из кадра, данный трекер может определить, когда слежение не удалось. Это является одним из его основных преимуществ.

Увеличение количества объектов, обнаруженных автоматической системой, позволяет определить различные ситуации на уровнях наглядно-действенного и наглядно-образного мышления. Чтобы перейти от простого распознавания объектов, основывающегося на восприятии (в данном случае, визуальном), к распознаванию через нахождение различных закономерностей для конкретного объекта, необходимо провести сложную и крайне трудоёмкую работу по определению различных закономерностей.

Наборы данных для обучения сверточной нейросети могут быть получены из виртуальной реальности. Тестирование проводится в реальной среде, использующей визуальную запись информации.

Результаты работы

После определения архитектуры и разработки алгоритма для сверточной нейронной сети требуется провести проверку корректности созданного программного обеспечения. Изначально следует убедиться в том, что дефицит обучающих данных является критической сложностью в процессе обучения нейронных сетей. Для подтверждения этого предположения, нейронная сеть обучалась определению чайника на небольшой объеме информации. В ходе указанного эксперимента было получено 10 изображений чайника, при этом минимальный требуемый объем составляет 100 изображений.

После обучения нейронной сети были переданы фотографии чайников из другого набора, отличного от того, на котором была обучена сеть. Из ста фотографий нейронная сеть, обученная на недостаточном количестве тестовых данных, правильно идентифицировала испытуемого только 20 раз, что, как легко подсчитать, составляет лишь 20% успешных результатов теста.

После получения результата, который будет использоваться для сравнения работы предложенной в статье программы, требуется создать простой и удобный интерфейс, демонстрирующий работу программы. Необходимы тесты не только для подтверждения корректного функционирования нейронной сети, но также для случаев представления ошибочных данных для анализа и объектов, неизвестных программе. Для достижения этих целей было создано программное обеспечение для тестирования способности определять объект на фотографии. Этот программный продукт обладает простым оконным интерфейсом (рис. 4), который требует от пользователя указания пути к файлу с изображением объекта для распознавания. После нажатия кнопки подтверждения программа должна отобразить сообщение с наименованием объекта, который нейронная сеть определила на фотографии.

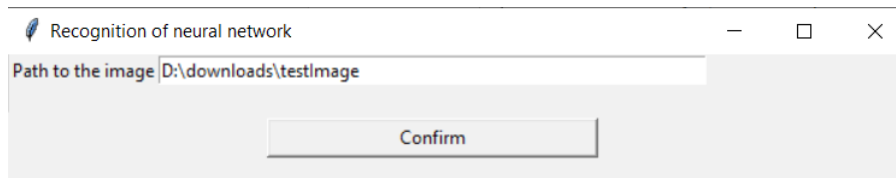


Рисунок 4 - Интерфейс оконного приложения для проверки работы созданной программы

Если вы введете неверный путь к файлу, отобразится сообщение о том, что указан неверный путь к изображению. Такая реакция возникает, когда вы вводите путь, по которому файл не существует, или если по указанному адресу находится файл с расширением, отличным от расширения файла изображения.

Если на вход программы поступает изображение, на котором не изображен чайник, то, как и ожидалось, программа не может распознать этот объект и выдает сообщение об этом.

После того как все возможные варианты неудачного исхода работы программы будут проверены, можно протестировать реакцию нейронной сети на картинку с изображением чайника.

Для теста на распознавание была сделана фотография настоящего чайника (рис.5), которая выглядит как трехмерная модель чайника, на которой была обучена нейронная сеть.



Рисунок 5 - Фотография чайника, на котором проводилось тестирование

После запуска программы становится ясно, что нейронная сеть правильно распознала, что на введенном изображении находится чайник.

После первоначального успешного тестирования было выбрано другое изображение чайника (рис. 6а), которое немного отличается от первого тестового изображения и модели, на которой была обучена программа. Сделанная фотография отличается от первой наличием золотой окантовки на верхних частях чайника и его крышке.

После запуска программы был получен ответ о том, что нейронная сеть распознала на входном изображении чайник. Данный тест показывает, что наличие некоторых дополнительных цветов на тестируемом изображении не влияет на корректность вывода программы. Для третьего теста была выбрана фотография чайника другого цвета с нанесенными на него узорами (рис. 6б).



Рисунок 6 - Образцы объектов на изображениях, использованных для второго и третьего тестов:
 а - чайник, который по дизайну очень похож на те, что использовались в обучающем наборе; б - чайник, который немного отличается от тех, что использовались в обучающем наборе (имеет сложные цветочные узоры)

Программа корректно обработала введенное изображение, определив на нем чайник. Испытание прошло успешно и продемонстрировало, что обученная нейросеть способна узнавать чайники не только белого цвета, но и при незначительных отличиях от изображений, использованных при обучении ИИ. Созданное ПО способно определять объекты разных цветов с разнообразными дополнительными изображениями на них. Четвертое тестирование проводилось на основе фотографии чайника, заметно отличающегося от обучающего объекта (рис. 7а). Помимо наличия рукоятки, он также имеет другую форму и укороченный носик. Несмотря на очевидные различия тестируемого объекта, программа успешно определила его как чайник. Для заключительного демонстрационного испытания была выбрана фотография кружки (рис. 7б). Такой выбор сделан благодаря тому, что кружка имеет схожий смысл с чайником, однако отличается функционально и визуально.



Рисунок 7 - Образцы предметов на изображениях, использованных для четвертого и пятого тестов:
 а - чайник, существенно отличающийся от обучающего набора; б - необычная чашка с частичным визуальным сходством с чайником

На введенной фотографии программа выдала ошибку, что не смогла распознать объект, который изображен. Данный тест показывает, что обученная нейронная сеть корректно распознает только те объекты, на которых она была натренирована. Таким образом, можно воспринимать этот тест как успешно пройденный.

Таблица 1 - Результаты тестирований

Тип объекта	Количество тестов	Точность
Чайник, похожий на тренировочный набор (как на рис. 6а)	100	93
Чайник, который немного отличается от тренировочного набора (как на рис. 6б)	100	87
Чайник, отличный от тренировочного набора (как на рис. 7б)	100	79
Другой предмет посуды (как на рис. 7а)	100	99

Далее было проведено расширенное тестирование нейронной сети, в рамках которого в программу распознавания были отправлены изображения чайников, схожие с обучающими, слегка отличающиеся и очень разные (табл. 1). По 100 фотографий было отобрано для каждой категории тестируемых предметов, чтобы грубо оценить процент успеха

программы. Критерием для отбора каждой группы служило качество изображения: выбирались фотографии, на которых предмет занимает большую часть кадра. Нужно пояснить, что этот набор данных не должен рассматриваться как результат исчерпывающего или подробного тестирования; он лишь служит для приблизительной оценки работы программы, ибо готового решения для столь всеобъемлющего тестирования не существует.

Также было проведено испытание, в ходе которого использовались другие кухонные предметы, для которых программа не была подготовлена. В этом случае успешным считается тест, в котором нейросеть не определила чайник.

После испытания программы видно, что разработанная и подготовленная нейронная сеть корректно распознает чайник на изображениях, что говорит о том, что такой метод обучения искусственной сети на объемных моделях изображений является эффективным.

Конечно, если на изображении присутствует чайник, но иной формы, нейронная сеть хуже определит принадлежность данного предмета. Более того, чем значительнее различия, тем выше вероятность некорректного определения предмета. Эту проблему можно решить, увеличив разнообразие форм предметов в различных обучающих наборах данных.

Заключение

Поскольку эффективность разработанной и обученной сверточной нейронной сети в распознавании 3D-моделей была экспериментально доказана, ее последующее развитие позволит получать результаты обработки не только по одной 3D-модели, но и по нескольким 3D-моделям, расположенным на одном изображении. Увеличение разнообразия возможных форм и свойств обучающих данных позволит улучшить корректность распознавания объектов, даже если они значительно отличаются от обычного представления обучающего объекта.

Затем, на основании определяемых объектов, возможно установление различных ситуаций посредством образного мышления на программном уровне. Как упоминалось ранее, чтобы перейти от простого распознавания объектов, основывающегося на восприятии (в данном случае, визуальном), к распознаванию посредством нахождения разнообразных закономерностей, требуются сложные и весьма продолжительные работы по выявлению разнообразных закономерностей, свойственных определенному объекту. Нейросети отлично подходят для выполнения таких задач. При корректно структурированном обучении нейросети время на это обучение существенно сокращается по сравнению со временем, которое требуется человеку и компьютеру для определения набора атрибутов, связанных с конкретным объектом.

В дальнейшем планируется разработка такой системы искусственного интеллекта, которая будет способна не только распознавать все объекты на изображении, но и устанавливать ситуацию, происходящую на снимке. К примеру, обучить систему распознаванию не только чайника, но и чашек, находящихся рядом с ним. После корректного распознавания программа сможет сделать вывод, что в этой ситуации содержимое чайника может быть разлито по чашкам.

Конфликт интересов

Не указан.

Рецензия

Все статьи проходят рецензирование. Но рецензент или автор статьи предпочли не публиковать рецензию к этой статье в открытом доступе. Рецензия может быть предоставлена компетентным органам по запросу.

Conflict of Interest

None declared.

Review

All articles are peer-reviewed. But the reviewer or the author of the article chose not to publish a review of this article in the public domain. The review can be provided to the competent authorities upon request.

Список литературы / References

1. Shimonishi H. Energy Optimization of Distributed Video Processing System using Genetic Algorithm with Bayesian Attractor Model / H. Shimonishi, M. Murata, G. Hasegawa [et al.] // Proc. of 2023 IEEE 9th International Conference on Network Softwarization (NetSoft). — 2023. — DOI: 10.1109/NetSoft57336.2023.10175483
2. Cha S.-H. Comprehensive survey on distance/similarity measures between probability density functions / S.-H. Cha // International Journal of Mathematical Models and Methods in Applied Sciences. — 2007. — 1(4). — P. 300-307.
3. Hamidreza Kasaei S. OrthographicNet: A Deep Transfer Learning Approach for 3D Object Recognition in Open-Ended Domains / S. Hamidreza Kasaei // IEEE/ASME Transactions on Mechatronics. — 2021. — 26(6). — P. 2910-2921. — DOI: 10.1109/TMECH.2020.3048433
4. Golovko V. Pattern Recognition and Information Processing. PRIP 2019 / V. Golovko, A. Kroshchanka, E. Mikhno et al. // Communications in Computer and Information Science. — 2019. — Vol 1055. — DOI: 10.1007/978-3-030-35430-5_4.
5. Anto Bennet M. Performance Evaluation of Video Surveillance Using Mete, Melt and Nidc Technique / M. Anto Bennet, R. Srinath, D. Abirami [et al.] // International Journal on Smart Sensing and Intelligent Systems. — 2022. — 10(5). — P. 25-54. — DOI: /10.21307/ijssis-2017-234
6. Lukezic A. Discriminative correlation filter tracker with channel and spatial reliability / A. Lukezic, T. Vojr, L.C. Zajc [et al.] // International Journal of Computer Vision. — 2018. — 126(7). — P. 671-688.
7. Kochanov A. Digital processing of satellite images using neural network algorithms / A. Kochanov, V. Zolotukhin, V. Mironenko [et al.] // Journal of Physics: Conference Series. — 2022. — 2373 — P. 062026.
8. Сорокин А.Б. Сверточные нейронные сети: примеры реализаций / А.Б. Сорокин, Л.М. Железняк, Е.А. Зикеева. — М.: РТУ МИРЭА, 2020. — 158 с.
9. Николенко С. Глубокое обучение. Погружение в мир нейронных сетей / С. Николенко, А. Кудрин, Е. Архангельская. — СПб.: Питер, 2018. — 480 с.
10. Шолле Ф. Глубокое обучение на Python / Ф. Шолле. — СПб.: Питер, 2018. — 400 с.
11. Никулин Е.А. Компьютерная графика. Модели и алгоритмы / Е.А. Никулин. — СПб: Лань, 2017. — 708 с.

Список литературы на английском языке / References in English

1. Shimonishi H. Energy Optimization of Distributed Video Processing System using Genetic Algorithm with Bayesian Attractor Model / H. Shimonishi, M. Murata, G. Hasegawa [et al.] // Proc. of 2023 IEEE 9th International Conference on Network Softwarization (NetSoft). — 2023. — DOI: 10.1109/NetSoft57336.2023.10175483
2. Cha S.-H. Comprehensive survey on distance/similarity measures between probability density functions / S.-H. Cha // International Journal of Mathematical Models and Methods in Applied Sciences. — 2007. — 1(4). — P. 300-307.
3. Hamidreza Kasaei S. OrthographicNet: A Deep Transfer Learning Approach for 3D Object Recognition in Open-Ended Domains / S. Hamidreza Kasaei // IEEE/ASME Transactions on Mechatronics. — 2021. — 26(6). — P. 2910-2921. — DOI: 10.1109/TMECH.2020.3048433
4. Golovko V. Pattern Recognition and Information Processing. PRIP 2019 / V. Golovko, A. Kroshchanka, E. Mikhno et al. // Communications in Computer and Information Science. — 2019. — Vol 1055. — DOI: 10.1007/978-3-030-35430-5_4.
5. Anto Bennet M. Performance Evaluation of Video Surveillance Using Mete, Melt and Nidc Technique / M. Anto Bennet, R. Srinath, D. Abirami [et al.] // International Journal on Smart Sensing and Intelligent Systems. — 2022. — 10(5). — P. 25-54. — DOI: /10.21307/ijssis-2017-234
6. Lukezic A. Discriminative correlation filter tracker with channel and spatial reliability / A. Lukezic, T. Vojr, L.C. Zajc [et al.] // International Journal of Computer Vision. — 2018. — 126(7). — P. 671-688.
7. Kochanov A. Digital processing of satellite images using neural network algorithms / A. Kochanov, V. Zolotukhin, V. Mironenko [et al.] // Journal of Physics: Conference Series. — 2022. — 2373 — P. 062026.
8. Sorokin A.B. Svertochnye nejronnye seti: primery realizacij [Convolved neural networks: examples of implementations] / A.B. Sorokin, L.M. Zheleznyak, E.A. Zikeeva. — M.: RTU MIRJeA, 2020. — 158 p. [in Russian]
9. Nikolenko S. Glubokoe obuchenie. Pogruzenie v mir nejronnyh setej [Deep Learning. Diving into the world of neural networks] / S. Nikolenko, A. Kudrin, E. Arhangel'skaja. — SPb.: Piter, 2018. — 480 p. [in Russian]
10. Sholle F. Glubokoe obuchenie na Python [Deep learning in Python] / F. Sholle. — SPb.: Piter, 2018. — 400 p. [in Russian]
11. Nikulin E.A. Komp'yuternaja grafika. Modeli i algoritmy [Computer Graphics. Models and algorithms] / E.A. Nikulin. — SPb: Lan', 2017. — 708 p. [in Russian]