Как известно, самая перспективная и популярная задача нейронных сетей – технология распознавания объектов. Они либо по-отдельности, либо в интегрированном виде используются в таких сферах, как безопасность, видеонаблюдение, сканирование и создание изображений, маркетинг и реклама, дополненная реальность и поиск изображений.

В настоящее время создаются и уже используются нейронные сети, в которых машины способны распознавать знаки на бумажном носителе и банковских пластиковых картах, электронные подписи на официальных документах и так далее. Эти функции упрощают труд человека и повышают точность и надежность различных рабочих процедур. Но обучить компьютерную технику распознавать объекты не так уж и легко. Одна из трудностей заключается в том, что компьютер видит не так же, как и человек. У компьютера нет жизненного опыта и способности так же, как человеческий мозг определять объекты на изображения и видео. Изначально 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. Выделив признаки и их комбинации для идентификации похожих объектов, можно научить модель машинного обучения распознавать нужные типы закономерностей. Конечно, даже после загрузки нескольких исходных данных модели могут неверно распознавать некоторые объекты.

Допустим, если рассматривать сферу видеонаблюдения, то ее основой является анализ, первой фазой которого будет распознавание объекта. Затем искусственный интеллект с помощью машинного обучения распознает действия и классифицирует их. Но для того, чтобы распознать изображение, нейросеть должна быть прежде обучена на исходных данных. Это очень похоже на нейронные связи в человеческом мозге — мы обладаем определенными знаниями, получаемыми в течение жизни, видим объект, анализируем его и идентифицируем. Также нейронные сети очень требовательны к размеру и качеству дата-сета, на котором она будет обучаться. Дата-сет можно загрузить из открытых источников или собрать самостоятельно. На практике это означает, что до определённого предела, чем больше скрытых слоев в нейронной сети, тем точнее будет распознан объект. Как это реализуется: картинка разбивается на маленькие участки, вплоть до нескольких пикселей, каждый из которых будет входным нейроном. С помощью синапсов сигналы передаются от одного слоя к другому. Во время этого процесса сотни тысяч нейронов с миллионами параметров сравнивают полученные сигналы с уже обработанными данными.

Другими словами, если мы просим машину распознать фотографию кошки, мы разобьем фото на маленькие кусочки и будем сравнивать эти слои с миллионами уже имеющихся изображений кошек, значения признаков которых сеть выучила.

Распознавание объектов — важная задача компьютерного зрения, используемая для обнаружения экземпляров визуальных объектов определенных классов в цифровых изображениях, таких как фотографии или видеозаписи. Целью обнаружения объектов является разработка вычислительных моделей, которые предоставляют наиболее фундаментальную информацию, необходимую приложениям компьютерного зрения: «Какие объекты и где они находятся?».

Распознавание объектов может выполняться с использованием либо традиционных методов обработки изображений, либо современных сетей глубокого обучения.

Методы обработки изображений, как правило, не требуют исходных данных для обучения и по своей природе неконтролируемы.

Достоинства: следовательно, эти задачи не требуют аннотированных изображений, где люди маркировали данные вручную.

Недостатки: эти методы ограничены несколькими факторами, такими как сложные сценарии, окклюзия, освещение и тени, и эффект беспорядка.

Методы глубокого обучения обычно зависят от контролируемого или неконтролируемого обучения, при этом контролируемые методы являются стандартом в задачах компьютерного зрения. Производительность ограничена вычислительной мощностью графических процессоров, которая стремительно растет с каждым годом.

Достоинства: Обнаружение объектов с помощью глубокого обучения значительно более устойчиво к окклюзии, сложным сценам и сложному освещению.

Недостатки: требуется огромное количество обучающих данных; процесс аннотации изображений является трудоемким и дорогостоящим.

Сегодня обнаружение объектов глубокого обучения широко признано исследователями и используется компаниями, занимающимися компьютерным зрением, для создания коммерческих продуктов.