



# ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ЕГО ВЛИЯНИЕ НА МАШИННОЕ ЗРЕНИЕ

ВИНН ХАРДИН (WINN HARDIN)

ПЕРЕВОД И ДОПОЛНЕНИЯ: ВЛАДИМИР РЕНТЮК

Многие функциональные возможности для развития человеческого мозга можно использовать при разработке оборудования. Внедрение искусственного интеллекта в системы машинного зрения и применение более современных методов обучения позволят машине думать почти так же, как человеку.

Когда в 1950-х гг. появилось понятие «думающая машина» (вычислительная машина с элементами искусственного интеллекта), возникла и большая озабоченность касательно возможностей и путей развития этой новой области. С тех пор в массовой культуре часто разыгрываются сцены восстания наделенных интеллектом машин. Начало было положено выходом культового фильма «Космическая одиссея 2001 года» Стенли Кубрика («2001: A Space Odyssey», Stanley Kubrick, 1968), а одной из последних кинокартин на эту тему стала «Из машины» Алекса Гарленда («Ex Machina», Alex Garland, 2014).

Несмотря на то, что искусственный интеллект (ИИ) еще не охватил все сферы нашего общества, достиг-

нутые улучшения в области хранения и обработки данных уже позволили разработать когнитивные системы<sup>1</sup>, такие как IBM Watson, которые призваны убрать интуитивные догадки из действий людей, принимающих важные решения. Однако в большинстве случаев ИИ применяют для выполнения гораздо более скромных задач, например, распознавания объектов и образов.

Использование ИИ в системах машинного зрения предполагает, что машина сможет принимать более сложные решения, выходящие за пределы современных возможностей. Но возникает вопрос — готова ли сама технология к появлению таких систем в промышленности?

## ВОЗМОЖНОСТИ ИИ В СИСТЕМАХ МАШИННОГО ЗРЕНИЯ

Применение ИИ в системах машинного зрения требует глубокого обучения машин. В широком

понимании ИИ — это способность компьютера имитировать человеческий интеллект. Глубокое машинное обучение позволяет компьютерам действовать без явного программирования, т. е. они могут учиться на собственном опыте.

Благодаря нескольким знаковым событиям, произошедшим за последнее десятилетие, глубокое обучение систем машинного зрения из возможности превратилось в реальность. «Новые методы построения нейронных сетей, наличие достаточной вычислительной мощности в графических процессорах (англ. Graphics Processing Unit, GPU), а также возможность использования «больших данных» открыли нам путь к применению ИИ для обработки изображений (образов)», — говорит Оливье Деспонт (Olivier Despont), отвечающий за развитие бизнеса в компании ViDi Systems, швейцарского создателя программного обеспечения для глубокого обучения машин.

<sup>1</sup> Когнитивная система, или когнитивная структура (от лат. cognitio — познание), применительно к рассматриваемому в статье случаю — это система познания на основе ИИ, сложившаяся в результате его обучения и накопленного им опыта. На основе этой системы ставятся цели и принимаются решения о том, как надо действовать в той или иной ситуации, чтобы избежать когнитивного диссонанса. Ранее это понятие ассоциировалось исключительно с человеком и учитывало особенности его характера, воспитания и знания об окружающем мире. — Прим. пер.

Глубокое обучение является весьма многообещающей технологией по сравнению с традиционными подходами, используемыми в системах машинного зрения. В отличие от традиционных методов с применением программного обеспечения для обработки изображений, в новых системах используется подход, основанный на правилах. «ИИ — это следующий шаг, который мы принимаем, когда имеем дело с тем, что нелегко охарактеризовать с достаточной точностью и однозначностью, или с нелинейным процессом. В таких случаях для достижения следующего уровня повторяемости в принятии решений мы должны внедрить ИИ в машины, — считает Уоллес Латимер (Wallace Latimer), директор по продажам специализированных оптических систем в компании FISBA LLC. — В то время как линейные алгоритмы создают очень узкое пространство для маневра, сочетание ИИ с глубоким обучением расширяет границы применения, поскольку оно может предоставить нам намного больше вариаций. Благодаря этому можно усложнить оценку: осуществлять ее не только по принципу «что такое хорошо или что такое плохо», но и «почему это хорошо или плохо». С большими возможностями в части гибкости вы сможете сосредоточиться на том, что дает наиболее эффективный результат в принятии решений, и уменьшить число изменений по входным параметрам».

В настоящее время на рынке систем машинного зрения существует, по крайней мере, одна система с глубоким обучением — ViDi Suite от компании ViDi Systems. Эта система является первым коммерчески доступным программным обеспечением, которое выполнено на основе глубокого обучения, предназначено для анализа изображений и адаптировано для использования в промышленности. Программное обеспечение, которое интегрируется со стандартными библиотеками обработки изображений, в части «познания мира» действует так, как бы это делал ребенок.

«Вы не учитесь ребенка, используя подход, основанный на четких правилах, т. е. объясняя ему, что такое дом, — говорит Деспонт. — Основываясь всего лишь на нескольких примерах, наш мозг даже в раннем возрасте может получить информа-

В литературе проблема восстания машин поднималась намного раньше кино. Так, писатель-фантаст и философ Айзек Азимов в рассказе «Хоровод» (1942) сформировал три основных закона робототехники, в нашем случае — искусственного интеллекта. Эти правила, не утратившие своей актуальности, гласят:

1. Робот не может причинить вред человеку или своим бездействием допустить, чтобы человеку был причинен вред.
2. Робот должен повиноваться всем приказам, которые дает человек, кроме тех случаев, когда эти приказы противоречат Первому Закону.
3. Робот должен заботиться о своей безопасности в той мере, в которой это не противоречит Первому или Второму Законам.

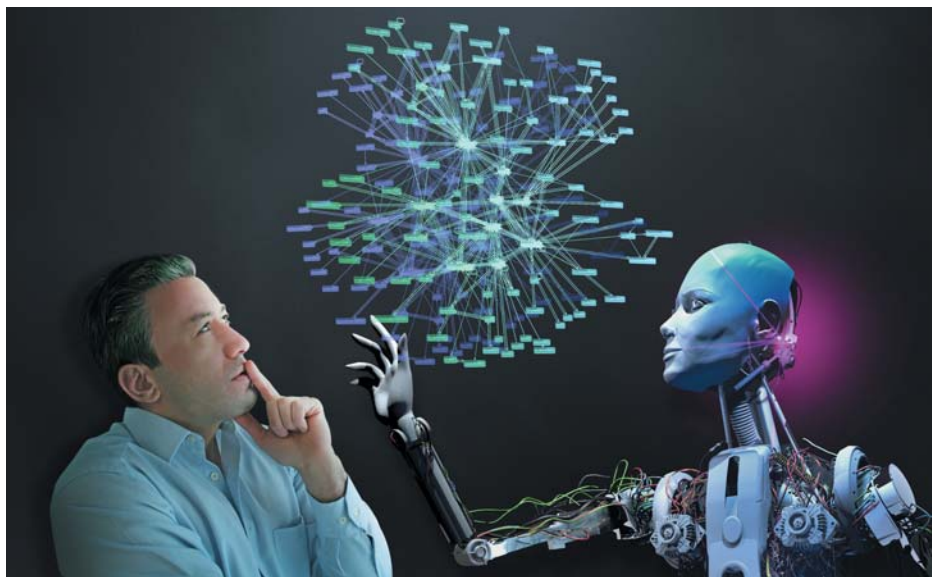
цию о том, что делает дом именно домом. Наша система работает так же, как и человеческий мозг».

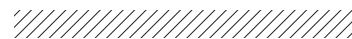
Программное обеспечение ViDi Suite состоит из трех инструментов. Первый, ViDi Blue, находит и обнаруживает отдельный объект или сразу несколько внутри общего изображения. Этот инструмент локализует и идентифицирует сложные и простые объекты, изучая их с помощью имеющихся в его распоряжении аннотированных изображений. Другой инструмент, ViDi Red, обнаруживает аномалии, анализируя нормальный внешний вид объекта, включая его вариации, а также сегментирует определенные области в изображениях. Наконец, третий из них, ViDi Green, используется для классификации объектов. Он обучается разделять объекты в различные классы на основе базы собранных и уже описанных изображений.

Еще одним важным преимуществом использования технологии глубокого

обучения по сравнению с традиционными решениями, применяемыми в системах машинного зрения, является то, что она может сократить время, необходимое для разработки программ машинного зрения. «При использовании подхода с классическим видением многие приложения нуждаются как минимум в двух месяцах разработки и отладки их программного обеспечения, — отмечает Деспонт. — Использование продуктов ViDi позволяет завершить всю разработку буквально за считанные часы».

В отличие от систем ИИ, которые используют серверные фермы как основу для своего программного обеспечения (например, как системы, разработанные компаниями Facebook, Google и IBM), системы компании ViDi обучаются с помощью одного высокопроизводительного графического процессора компании NVIDIA. Причем весь процесс, по словам Оливье Деспонта, занимает считанные минуты, а не дни или





месяцы, которые требуются для программирования и записи исходной информации в параметрической форме с помощью программного обеспечения IBM Watson.

«Вместо того, чтобы использовать миллионы или миллиарды изображений для обучения системы, мы рекомендуем начинать с 30-50 репрезентативных изображений высокого качества, — говорит Деспонт. — При этом для обработки информации или обучения мы не отправляем изображения в облачную серверную ферму. Благодаря этому можно запускать все на одном персональном компьютере с одним графическим процессором и сохранять интеллектуальные права собственности на свои изображения».

### ВОЗМОЖНОСТИ И ПРОБЛЕМЫ

Глубокое обучение позволяет использовать системы машинного зрения для гораздо более широкого спектра применений. «ИИ хорошо подходит, например, для анализа продуктов питания в общей массе, когда вы хотите проверить пончики или куски мяса, экземпляры которых существенно отличаются друг от друга», — отмечает Бруно Менар (Bruno Ménard), менеджер по программным продуктам в компании Teledyne Dalsa.

При этом процесс идентификации можно будет сделать более сложным. Для пояснения Менар приводит

в качестве примера традиционные приложения обнаружения дефектов. «Трудно запрограммировать компьютер с традиционными алгоритмами для определения дефекта без необходимости повторять настройки каждый раз, когда появляется новый дефект, — сказал он. — Но используя ИИ при большом количестве образцов, вы можете получить действительно хорошее определение того, что является качественной деталью или продуктом, а что браком».

Система машинного зрения с внедренным ИИ найдет применение в дополнительных задачах контроля и в итоге выйдет за привычные рамки промышленной автоматизации. По словам Латимера, глубокое обучение будет выгодно использовать на таких рынках, как: медицина, медико-биологические исследования, продукты питания, а также для проверки продукции на подлинность с целью выявления контрафакта и для сортировки пиломатериалов.

«Это как раз те отрасли, в которых имеется много серых пятен и условностей в принятии решений, — отмечает Латимер. — Например, как оценить, достаточно ли хорошее это яблоко? Трудно создать некое линейное правило, чтобы ответить на этот вопрос. Глубокое обучение позволит многим приложениям стать более эффективными и иметь высокую повторяемость в результате принятия ими решений».

Со своей стороны, Деспонт из компании ViDi Systems предполагает, что глубокое обучение будет востребовано и в других отраслях: медицинской диагностике, системах видеонаблюдения, автономных транспортных средствах, а также в «умном» сельском хозяйстве (для проверки или анализа карт). «ИИ — это наше будущее, и с помощью этой технологии мы достаточно быстро и эффективно поможем людям решать сложные задачи, поскольку вычислительные возможности процессоров удваиваются практически каждые полтора года», — говорит Деспонт.

Многие специалисты в области систем машинного зрения признают все преимущества и ясно понимают перспективы, которые ИИ и глубокое обучение дают индустрии распознавания образов, но считают, что полный потенциал ИИ не будет раскрыт и реализован в полной мере в течение еще как минимум трех-пяти лет. Более того, ИИ не обязательно будет единственным универсальным решением для всего того, что повлияет на традиционные системы машинного зрения и обработку изображений.

Менар также отмечает два основных недостатка в системах ИИ. «Во-первых, нужно много тренировок... и необходимо создать своеобразного эксперта для достижения следующего уровня классификации, — говорит он. — Второй недостаток проявляется, когда система, выполненная на базе ИИ, уже обучена, но классификация терпит неудачу. Эту проблему трудно решить. И у вас не будет другой попытки выполнить ее переподготовку с новым образцом».

Прежде чем ИИ станет обычным явлением для систем машинного зрения, как считают эксперты, индустрия должна позволить гораздо большему количеству игроков сделать первые непростые шаги для внедрения этой технологии. «Из нашего сегмента рыночной ниши мы будем наблюдать, как такие мировые лидеры, как, например, компания Google, с ее невероятными возможностями в части инвестиций, будут справляться с этой технологией, — подводит итог Латимер. — Наша отрасль промышленности еще не настолько богата, чтобы вкладывать время и деньги в необходимых объемах. Поэтому нам ничего не остается, кроме как использовать опыт больших игроков». ●

