

УДК 004.089

ПРИМЕНЕНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ГЕНЕРАЦИИ ИЗОБРАЖЕНИЙ

Иванов Д.А.

студент

*Поволжский Государственный Университет Телекоммуникаций и
Информатики
Самара, Россия*

Лиманова Н.И.

кандидат технических наук

*Поволжский Государственный Университет Телекоммуникаций и
Информатики
Самара, Россия*

Аннотация

Применение искусственного интеллекта для генерации изображений является актуальной темой, так как это может быть полезным во многих областях, включая графический дизайн, производство фильмов и видеоигр, медицину и многое другое. Генерирование уникальных эскизов с помощью ИИ может значительно сократить время на их создание и увеличить качество конечной продукции. Более того, это позволяет создавать образы, которые были бы очень сложно или невозможно сделать вручную. В отраслях, где требуется большое количество иллюстраций и изображений, это может привести к снижению затрат на персонал и сокращению временных затрат на производство. Кроме того, профессиональным сотрудникам может потребоваться диагностика магнитно-резонансной томографии или рентгеновских снимков в медицинских учреждениях, что значительно может быть ускорено с помощью использования искусственного интеллекта. В целом, применение искусственного интеллекта для генерации своеобразного искусства имеет огромный потенциал для технологического прогресса и может привести к усовершенствованию многих областей деятельности.

Ключевые слова: искусственный интеллект, нейронные сети, технология, изображение, продуктивность, ускорение процесса.

Дневник науки | www.dnevniknauki.ru | СМИ ЭЛ № ФС 77-68405 ISSN 2541-8327

USING ARTIFICIAL INTELLIGENCE TO GENERATE IMAGES

Ivanov D.A.

student

Povolzhskiy State University of Telecommunications and Informatics

Samara, Russia

Limanova N.I.

Candidate of Technical Sciences

Povolzhskiy State University of Telecommunications and Informatics

Samara, Russia

Abstract

The use of artificial intelligence for image generation is a hot topic, as it can be useful in many fields, including graphic design, film and video game production, medicine and more. Image generation using artificial intelligence can significantly reduce the time it takes to create images and increase the quality of the final product. What's more, it can create images that would be very difficult or impossible to do by hand. In industries that require a large number of illustrations and images, this can lead to lower staff costs and reduced production time. In addition, professional staff may need to diagnose MRIs or X-rays in medical facilities, which can be greatly accelerated by the use of artificial intelligence. Overall, the use of artificial intelligence to generate images has enormous potential for technological advancement and could lead to improvements in many fields of endeavour.

Keywords: artificial intelligence, neural networks, technology, image, productivity, process acceleration

Генерация изображений с использованием искусственного интеллекта — это процесс, при котором компьютерные алгоритмы создают картинки, не являющиеся копиями реальности. В этой технологии используется глубокое обучение, позволяющее компьютеру интуитивно понимать, что должно находиться на картинке [5].

В основе этой технологии лежит нейронная сеть, которая используется для определения того, как должно выглядеть изображение, и как его можно создать наиболее естественным способом. Обучение такой нейронной сети включает в себя тысячи иллюстраций, которые могут быть разных размеров, форм и цветов. Когда сеть обучена, ее можно использовать для создания новых, никогда не виденных ранее воспроизведений [4].

Одним из примеров приложения сотворения изображений, может быть, порождение игровых сред, где миры и персонажи должны быть созданы с нуля. Учитывая, что создание подобных изображений требует больших денежных и временных затрат, использование искусственного интеллекта для генерации изображений является намного более экономически выгодным решением [1].

Таким образом, создание изображений с использованием искусственного интеллекта является потенциально практичным и перспективным направлением исследований, которое может заменить затратные методы, использующие исполнительных художников и графических дизайнеров.

Представим основоположников данной технологии, а именно методы генерирования изображений, с которых и начался путь внедрения искусственного интеллекта в обучение моделей и соответственно генерацию уникальных изображений.

Генеративно-сопоставительные сети (GAN) — это нейронные сети, основанные на методе выработки искусственных данных на основе противоположных целей двух нейросетевых модулей - генератора и дискриминатора. Генератор создает синтетические данные путем преобразования случайных входных параметров в выходные данные, которые стараются походить на реальные данные. Дискриминатор, с другой стороны, старается различать эти синтетические данные от реальных.

При обучении GAN-ов, генератор и дискриминатор играют в игру, которую можно рассматривать как двухстороннюю игру Наше-Хоффера, где

каждый игрок пытается максимизировать свою выгоду, однако общий результат достигается только при равенстве выгод обоих игроков [2].

Использование GAN-ов для генерации изображений обладает многими преимуществами, такими как возможность создания новых изображений, которые находятся за пределами образов, используемых при обучении. Кроме того, данный нейросетевой подход позволяет генерировать качественные графические элементы с невероятной детализацией и естественностью, благодаря его дополнительным слоям нормализации и активации, таким как слой BN и активации ReLU.

Вариационные автокодировщики (VAE) — это класс глубоких генеративных моделей, которые позволяют изучать сложные стохастические процессы в пространстве скрытых переменных, обладая высокой гибкостью и мощным выразительным потенциалом. VAE основываются на идеи приближенного вывода, в котором задается вероятностная модель как суперпозиция нескольких слоев, каждый из которых функционирует как кодировщик и декодировщик [10].

В процессе сотворения картинок, VAE используются, чтобы найти примеры скрытых факторов, которые лежат в основе входных данных, и применяются для генерации новых уникальных иллюстраций путем генерации новых экземпляров скрытых переменных.

VAE улучшают производительность генерации путем применения дифференцируемой сэмпированной оценки нижней границы правдоподобия, для которой использовался синтез случайных шумов и их приписывание к каждому образцу генерируемых изображений. Этот процесс обеспечивает устойчивость при генерации изображений и позволяет построить эффективный метод генерации уникальных образов.

После того, как нейронные сети показали результаты, вдохновляющие на продолжение этой тематики, программные инженеры стали развивать эту

технологии, и каждая желающая компания стала делать свои программы искусственного интеллекта.

Одним из таких программных средств является NVIDIA StyleGAN, который является глубокой архитектурой генеративно-сопоставительных сетей (GAN), разработанная для генерирования высококачественных изображений из низкоразмерных векторов. Эта сеть стала популярной благодаря своему способу производства очень реалистичных графических составляющих [3].

Архитектура StyleGAN включает в себя множество слоев, каждый из которых выполняет определенную функцию, такую как конволюционные слои, блоки изменения масштаба и слои наращивания размерности.

Однако главным отличием StyleGAN является использование огромного количества дискретных нетривиальных улучшений в GAN-архитектуре, которые приводят к невероятно изящным и убедительным результатам генерации изображений.

Особое внимание в этой архитектуре уделяется тому, что на получении изображений должны работать многоуровневые модели, что увеличивает трудоемкость и сложность работы. Поэтому была использована техника декомпозиции, которая была более эффективной.

Одной из наиболее примечательных особенностей архитектуры StyleGAN является ее способность к управляемой генерации изображений. Это позволяет пользователю контролировать параметры создания, такие как освещение, цвет и форма, что приводит к более разнообразным и интересным результатам генерирования.

Кроме того, StyleGAN также обладает алгоритмом формирования изображений с высокой разрешающей способностью, который будет использоваться для создания видео и трехмерных моделей в будущем.

DALL-E 2 является продолжением достижений глубокого обучения в области генеративных моделей и имеет революционный потенциал в создании инновационных решений для автоматизированного создания изображений.

Архитектура DALL-E 2 основывается на принципе генеративного адверсариального обучения, используя глубокие нейронные сети и вычислительные технологии на базе графических процессоров [6].

Основной технологический элемент DALL-E 2 — это сеть состояний обратных связей, которая предоставляет информацию о том, какие комбинации входных данных порождают наилучший результат на выходе. Эта сеть работает вместе с вариационным автокодировщиком и глубокими сверточными нейронными сетями для генерации изображений с высокой степенью детализации и художественного качества.

Архитектура DALL-E 2 также использует механизмы самообучения, что позволяет ей обучаться на большом количестве данных и создавать бесконечно много новых и уникальных изображений. Это достигается за счет использования сложных алгоритмов, включающих в себя сверточные нейронные сети, оценочные функции и методы сравнения изображений. Эти механизмы способствуют созданию более точных и реалистичных изображений, которые соответствуют требованиям дизайнеров, художников и других креативных профессионалов.

В целом, DALL-E 2 является высокоэффективной и многосторонней системой, которая может создавать образы в различных стилях и форматах. Она имеет огромный потенциал для использования в различных приложениях, включая игры, анимацию, рекламу и научные исследования.

Существует множество способов использования реалистичных сгенерированных фотографий, и при этом, ограничений почти нет – от создания полноценных компьютерных игр и виртуальных путешествий, до разработки адаптивных интерфейсов и медицинских исследований. Конечно, искусственный интеллект и генерация изображений – технологии, которые еще на стадии развития, но перспективы светлые.

Когда говорят о применении искусственного интеллекта для создания изображений, обычно имеют в виду нейросети, способные сами «видеть» мир и Дневник науки | www.dnevniknauki.ru | СМИ ЭЛ № ФС 77-68405 ISSN 2541-8327

генерировать изображения, которые практически неотличимы от реальных. Это происходит благодаря глубокому обучению, при котором нейросеть проходит через многократные итерации, отличая правильные образцы от неправильных, и постепенно учась создавать собственные изображения на основе имеющейся информации [7].

Одно из основных применений этой технологии – это создание виртуальных миров. В компьютерных играх, виртуальных путешествиях, в разработке военных стратегий и многих других областях, нейросеть может создать необычайно реалистичные сцены, которые до этого можно было увидеть только в кино или на рисунках художников.

Еще одним примером использования этой методики является разработка адаптивных интерфейсов. Для людей с ограниченными возможностями зрения, слуха или моторики, создание альтернативных средств взаимодействия может быть очень трудным. С помощью нейросетей можно создавать адаптивные интерфейсы, которые будут адекватно реагировать на жесты и команды, даже если они не очень точны или различимы.

Наконец, генерация уникальных фотографий может быть полезна в медицине и науке. Медики могут использовать сгенерированные фото для тренировки на диагностике, а для ученых это может стать способом быстрого создания моделей для исследования новых лекарств или материалов.

Конечно, существует множество других областей применения, и это только начало. С развитием технологий и популяризации нейросетей, мы можем увидеть революцию во многих областях.

Сгенерированные реалистичные фотографии являются одним из возможных результатов применения искусственного интеллекта для генерации изображений. Эта технология использует алгоритмы машинного обучения, которые способны создавать фотографии, которые могут выглядеть такими, как будто они были сделаны настоящим фотоаппаратом.

Одним из основных преимуществ создания изображений является быстрота и эффективность процесса. Генерирование изображений может занимать меньше времени, чем создание реальных фотографий, а также позволяет получать большое количество изображений с меньшими затратами средств. Это особенно важно для сфер, где количественное состояние изображений оказывает влияние на повседневную жизнь людей, таких как медицина, наука и образование [8].

Однако для генерации картинок требуется большой объем данных для обучения. Чтобы создать качественные иллюстрации, необходимо иметь большой объем данных, таких как фотографии, из которых можно извлечь характеристики для обучающей выборки. Если данных недостаточно, соответственно, могут быть проблемы с качеством создаваемых абстракций.

Важным недостатком генерированных фотографий является ограничение на качество изображений. Хотя они могут выглядеть реалистично и качественно, они могут не соответствовать потребностям предприятий, особенно в сферах, связанных с высоким качеством, таких как реклама, дизайн, архитектура.

Также следует отметить, что эта технологическая система требует вычислительной мощности компьютера. Из-за сложности алгоритмов генерации, этот процесс требует мощных процессоров и графических карт. Это может быть дорого, особенно для компаний, которые хотят использовать создание изображений в своих бизнес-операциях [9].

Таким образом, использование различных моделей или программ, основанных на искусственном интеллекте и глубоком обучении, уже в ближайшем будущем будет помогать многим людям и организациям экономить время и ресурсы на воплощение своих мыслей в визуальное представление того, что может увидеть вместе с Вами нейронная сеть по заданным критериям и подсказкам, а то и переработки того, что Вам не особо нравится, в совершенно новое, получившее новую жизнь после взаимодействия с искусственным интеллектом.

Библиографический список

1. 3 Questions: How AI image generators work. URL: <https://www.csail.mit.edu/news/3-questions-how-ai-image-generators-work>
2. A Gentle Introduction to Generative Adversarial Networks (GANs). URL: <https://machinelearningmastery.com/what-are-generative-adversarial-networks-gans/>
3. A brief overview of NVIDIA StyleGAN. URL: <https://medium.com/imagescv/a-brief-overview-of-nvidia-stylegan-61cb24ec01f5>
4. AI Image Generators: How They Work and Why They Are Important. URL: <https://www.hypotenuse.ai/blog/ai-image-generator>.
5. Beginner's Guide to AI Image Generators. URL: <https://www.unite.ai/beginners-guide-to-ai-image-generators/>.
6. DALL-E – Wikipedia. URL: <https://en.wikipedia.org/wiki/DALL-E>
7. Generate Totally Realistic Images of This Person Does Not Exist with ImgCreator.ai. URL: <https://www.zmo.ai/generate-images-of-this-person-does-not-exist-with-ai/>
8. Pros and cons of AI-generated content. URL: <https://www.techtarget.com/whatis/feature/Pros-and-cons-of-AI-generated-content>
9. THE PROS AND CONS OF CREATING IMAGES USING AI TOOLS. URL: <https://www.apollotechnical.com/the-pros-and-cons-of-creating-images-using-ai-tools/3>
10. Understanding Variational Autoencoders (VAEs). URL: <https://towardsdatascience.com/understanding-variational-autoencoders-vaes-f70510919f73>

Оригинальность 89%