



Международный журнал информационных технологий и энергоэффективности

Сайт журнала:

<http://www.openaccessscience.ru/index.php/ijcse/>



УДК 004.8

ИСПОЛЬЗОВАНИЕ НЕЙРОННОЙ СЕТИ ДЛЯ ГЕНЕРАЦИИ ИЗОБРАЖЕНИЙ

Полежаева М.В., Кенжина Д.С., ¹Нерпин Е.С., Сафонова Т.В., ²Мокряк А.В.

ФГБОУ ВО "РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ ГИДРОМЕТЕОРОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ" Санкт-Петербург, Россия (192007, город Санкт-Петербург, Воронежская ул., д. 79) e-mail: ¹nerpin.evgeniy@gmail.com

²ФГБОУ ВО "САНКТ-ПЕТЕРБУРГСКИЙ УНИВЕРСИТЕТ ГОСУДАРСТВЕННОЙ ПРОТИВОПОЖАРНОЙ СЛУЖБЫ МИНИСТЕРСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ ПО ДЕЛАМ ГРАЖДАНСКОЙ ОБОРОНЫ, ЧРЕЗВЫЧАЙНЫМ СИТУАЦИЯМ И ЛИКВИДАЦИИ ПОСЛЕДСТВИЙ СТИХИЙНЫХ БЕДСТВИЙ ИМЕНИ ГЕРОЯ РОССИЙСКОЙ ФЕДЕРАЦИИ ГЕНЕРАЛА АРМИИ Е.Н.ЗИНИЧЕВА", Санкт-Петербург, Россия (196105, г. Санкт-Петербург, Московский проспект, д.149), e-mail: mokryakanna@mail.ru

В данной статье рассматриваются нейросети для генерации изображений, их основные принципы работы, методы обучения, примеры применения, а также этические и социальные аспекты. Описаны ключевые архитектуры и алгоритмы, такие как GAN, VAE и Diffusion Models, и их применение в различных областях, включая искусство, дизайн, медицину и маркетинг. Также обсуждаются проблемы авторских прав, создание поддельных изображений (deepfake), влияние на рынок труда и предвзятость данных. Статья подчеркивает необходимость разработки нормативных и правовых рамок для этичного и безопасного использования данных технологий.

Ключевые слова: Нейросети, генерация изображений, глубокое обучение, GAN, VAE, Diffusion Models, искусство, дизайн, медицина, маркетинг, этика, социальные аспекты, авторские права, deepfake, предвзятость данных

USING A NEURAL NETWORK TO GENERATE IMAGES

Polezhaeva M.V., Kenzhina D.S., ¹Nerpin E.S., Safonova T.V., ²Mokryak A.V.

RUSSIAN STATE HYDROMETEOROLOGICAL UNIVERSITY, St. Petersburg, Russia (192007, St. Petersburg, Voronezhskaya str., 79), e-mail: ¹nerpin.evgeniy@gmail.com

²ST. PETERSBURG UNIVERSITY OF THE STATE FIRE SERVICE OF THE MINISTRY OF THE RUSSIAN FEDERATION FOR CIVIL DEFENSE, EMERGENCIES AND ELIMINATION OF CONSEQUENCES OF NATURAL DISASTERS NAMED AFTER THE HERO OF THE RUSSIAN FEDERATION, GENERAL OF THE ARMY E.N. ZINICHEV, St. Petersburg, Russia (196105, St. Petersburg, Moskovsky prospekt, 149), e-mail: ¹mokryakanna@mail.ru

This article discusses neural networks for image generation, their basic principles of operation, teaching methods, application examples, as well as ethical and social aspects. Key architectures and algorithms such as GAN, VAE, and Diffusion Models are described, and their applications in various fields, including art, design, medicine, and marketing. Copyright issues, the creation of fake images (deepfake), the impact on the labor market and data bias are also discussed. The article emphasizes the need to develop a regulatory and legal framework for the ethical and safe use of these technologies.

Keywords: neural networks, image generation, deep learning, GAN, VAE, Diffusion Models, art, design, medicine, marketing, ethics, social aspects, copyright, deepfake, data bias.

Введение

За последние годы нейронные сети для генерации изображений стали одной из самых обсуждаемых и быстроразвивающихся областей в искусственном интеллекте. Благодаря достижениям в области глубокого обучения, современные алгоритмы могут создавать изображения, которые поразительно похожи на реальные. Эти технологии открывают новые возможности для креативных индустрий, таких как искусство, дизайн, кино и реклама, позволяя создавать оригинальные произведения, которые ранее были немислимы без участия человека.

Основная идея заключается в том, что нейросеть обучается на огромных массивах данных — изображениях, которые она анализирует и на основе которых она учится создавать новые. Эти нейросети, используя сложные архитектуры и алгоритмы, способны «понимать» и воспроизводить различные аспекты изображений: от базовых форм и цветов до сложных текстур и деталей.

Эти инновации привлекают внимание не только исследователей и разработчиков, но и широкой общественности, вызывая как восхищение, так и вопросы. Важно понимать, как именно работают эти технологии, какие алгоритмы лежат в их основе и какие перспективы они открывают для будущего.

В этой статье мы рассмотрим основные принципы работы нейросетей для генерации изображений и узнаем о наиболее популярных архитектурах и примерах их применения.

Функциональные особенности работы нейронных сетей

Нейросети, или нейронные сети, представляют собой вычислительные модели, вдохновленные структурой и функционированием мозга. Они состоят из множества взаимосвязанных узлов, называемых нейронами, которые работают совместно для обработки и анализа информации. Каждый нейрон принимает входные сигналы, обрабатывает их и передает результаты на следующие нейроны. Нейронные сети способны решать широкий спектр задач, от классификации данных до создания сложных изображений [1].

Основные принципы глубокого обучения

Глубокое обучение — это подмножество машинного обучения, которое использует многослойные нейронные сети для анализа данных. Основное отличие глубокого обучения от традиционных методов заключается в использовании глубоких архитектур, состоящих из большого числа слоев (иногда сотен или даже тысяч). Эти слои могут обнаруживать иерархические представления данных, начиная от простых признаков (например, краев на изображениях) и заканчивая сложными структурами (например, лицами или объектами).

Ключевые концепции глубокого обучения:

1. Слои нейронной сети:

- входной слой: принимает исходные данные (например, изображение).
- скрытые слои: обрабатывают данные и извлекают признаки. Могут быть как простыми (полносвязные слои), так и специализированными (сверточные слои для обработки изображений).
- выходной слой: выдаёт результат работы сети (например, классификация или сгенерированное изображение) [2].

2. Обучение:

- обучающие данные: набор данных, на которых сеть обучается. В случае генерации изображений это могут быть тысячи и миллионы изображений.
- функция потерь: мера ошибки сети, показывающая, насколько её вывод отличается от ожидаемого результата. Например, для задачи генерации изображений это может быть разница между реальным и сгенерированным изображением.
- оптимизация: процесс настройки весов нейронов для минимизации функции потерь. Обычно используется алгоритм обратного распространения ошибки в сочетании с методами оптимизации, такими как градиентный спуск.

3. Гиперпараметры: параметры модели, которые не обучаются непосредственно, а настраиваются перед началом обучения. Примеры гиперпараметров включают скорость обучения, количество слоев и нейронов в каждом слое [1].

Преимущества глубокого обучения

Глубокое обучение позволяет нейросетям автоматически извлекать признаки и шаблоны из данных, минимизируя необходимость ручного выбора признаков. Это делает его особенно мощным инструментом для обработки больших и сложных наборов данных, таких как изображения, аудио и текст. Благодаря способности обучаться на огромных объемах данных и адаптироваться к сложным задачам, глубокое обучение стало основой для многих современных достижений в области искусственного интеллекта, включая генерацию изображений.

Обучение нейронных сетей для генерации изображений

- Сбор и подготовка данных

Сбор и подготовка данных являются первым и ключевым этапом в обучении нейросети для генерации изображений, поскольку качество и разнообразие данных напрямую влияют на результаты. Источники данных могут быть разнообразными: изображения можно собирать из интернета, использовать базы данных изображений, такие как ImageNet и COCO, или создавать вручную. Для успешного обучения нейросетей требуются большие объемы данных, обычно в тысячах или даже миллионах изображений.

После сбора данных необходимо их подготовить. Один из важных методов подготовки — аугментация данных, которая увеличивает разнообразие обучающих данных и улучшает обобщающую способность модели. Аугментация включает такие методы, как повороты, сдвиги, масштабирование, изменение яркости и контрастности изображений. В некоторых случаях требуется разметка данных, чтобы нейросеть могла лучше понять структуру и содержимое изображений. Это особенно важно для задач сегментации или распознавания объектов. Нормализация данных также является важным шагом, включающим преобразование данных в единый формат и масштаб, например, изменение размера изображений и нормализацию пиксельных значений, что облегчает их обработку нейросетью [2].

- Алгоритмы и модели глубокого обучения

Для генерации изображений используются несколько ключевых архитектур и алгоритмов глубокого обучения.

Generative Adversarial Networks (GANs) состоят из двух нейросетей — генератора и дискриминатора. Генератор создает изображения, а дискриминатор оценивает их качество. В процессе обучения генератор стремится создавать изображения, неотличимые от реальных, в

то время как дискриминатор учится отличать реальные изображения от сгенерированных. Обе сети обучаются одновременно, соревнуясь друг с другом. Существует множество вариантов GAN, включая Conditional GAN (cGAN), StyleGAN и CycleGAN, каждый из которых решает специфические задачи генерации изображений [3].

Variational Autoencoders (VAEs) также состоят из двух частей — энкодера и декодера. Энкодер сжимает изображение в латентное пространство, а декодер восстанавливает его из этого латентного представления. Модель обучается минимизировать разницу между входным изображением и восстановленным, а также регуляризовать латентное пространство для обеспечения плавного и непрерывного представления данных. VAEs отлично подходят для генерации изображений, поскольку эффективно моделируют сложные распределения данных и обеспечивают плавные переходы между сгенерированными изображениями [4].

Diffusion Models генерируют изображения путем пошагового удаления шума из случайного шума с использованием обученных нейросетей, которые предсказывают, как должны выглядеть изображения на каждом шаге. Модель обучается добавлять шум к изображениям и затем восстанавливать их из зашумленных версий. Одним из популярных примеров является модель Denoising Diffusion Probabilistic Models (DDPM) [5].

Обучающие задачи для этих моделей включают генерацию реалистичных изображений и создание стилизованных изображений. В случае обучения GAN на большом наборе реальных изображений, цель состоит в генерации новых, фотореалистичных изображений. VAEs также могут быть обучены на изображениях для создания нового контента, соответствующего общей структуре и стилю исходных данных. Для создания стилизованных изображений можно использовать StyleGAN, который обучается на изображениях определенного стиля, например, картин известных художников, и создает новые изображения в этом стиле. CycleGAN, в свою очередь, обучается преобразовывать изображения из одного стиля в другой, например, из фотографий в картины.

Обучение нейросетей для генерации изображений включает в себя несколько ключевых этапов: сбор и подготовку данных, выбор и настройку алгоритмов и моделей глубокого обучения, а также проведение множества экспериментов для достижения наилучших результатов. Благодаря современным достижениям в области глубокого обучения, нейросети способны создавать изображения, которые поразительно реалистичны и разнообразны, открывая новые горизонты для творчества и инноваций.

Примеры использования

Нейронные сети для генерации изображений находят широкое применение в различных сферах, от искусства и дизайна до медицины и развлечений. В мире цифрового искусства нейросети, такие как GAN, позволяют художникам создавать уникальные произведения, которые поражают своей оригинальностью и глубиной. Например, приложения вроде DeepArt позволяют пользователям превращать свои фотографии в произведения, выполненные в стиле известных художников, что открывает новые горизонты для самовыражения и креативности [6].

В области дизайна и моды генеративные модели помогают создавать новые узоры и модели одежды. Компании, такие как NVIDIA, активно используют GAN для генерации инновационных дизайнов одежды и аксессуаров, что значительно ускоряет процесс разработки и позволяет предлагать потребителям уникальные продукты. Аналогично, в

архитектуре нейросети применяются для создания новых проектов зданий и интерьеров, предлагая нестандартные решения и инновационные подходы, которые иначе могли бы быть упущены [7].

Генерация фотореалистичных изображений — еще одно важное направление. С помощью технологий, таких как StyleGAN, можно создавать изображения людей, которые не существуют в реальности, что находит применение в маркетинге, киноиндустрии и других областях. Кроме того, нейросети способны восстанавливать поврежденные фотографии, заполняя утраченные части с высоким уровнем реализма, что особенно ценно в реставрации исторических изображений [7].

Не менее впечатляющими являются возможности нейросетей в создании видео. Модели, такие как Vid2Vid, позволяют генерировать реалистичные видеоролики на основе заданных условий, что может быть полезно в анимации и кино. Технологии deepfake позволяют заменять лица в видео, что используется как в развлекательных целях, так и в негативных, таких как распространение дезинформации [7].

В игровой индустрии и виртуальной реальности нейросети играют важную роль в создании контента. Они могут автоматически генерировать ландшафты, текстуры и объекты для игровых миров, что значительно сокращает время и усилия, необходимые для их разработки. Виртуальные персонажи и аватары могут быть сгенерированы на основе предпочтений пользователей, создавая уникальный и персонализированный опыт.

Медицина также выигрывает от применения генеративных нейросетей. В области медицинской визуализации нейросети улучшают качество изображений, таких как МРТ и КТ, что помогает в ранней диагностике заболеваний. Кроме того, генеративные модели помогают создавать синтетические данные для обучения других медицинских алгоритмов, что повышает точность диагностики и лечения.

В маркетинге и рекламе нейросети позволяют создавать оригинальные рекламные изображения и видео, соответствующие брендовым стандартам. Они также могут анализировать эффективность рекламных кампаний и предлагать оптимальные визуальные и текстовые решения для привлечения клиентов. Генерация A/B тестов с различными вариантами контента помогает определить наилучшие стратегии для достижения маркетинговых целей.

Таким образом, генеративные нейросети открывают множество возможностей для инноваций в самых разных областях. Эти технологии не только ускоряют и упрощают процессы создания контента, но и позволяют достигать результатов, которые ранее были невозможны, что делает их незаменимым инструментом в современном мире.

Этические и социальные аспекты

Использование нейросетей для генерации изображений приносит не только значительные выгоды, но и вызывает множество этических и социальных вопросов. Одной из наиболее острых проблем является вопрос авторских прав. Поскольку нейросети обучаются на огромных наборах данных, включающих защищенные авторским правом изображения, возникает вопрос о законности использования этих данных. Нередко для обучения нейросетей используются изображения без разрешения их владельцев, что может нарушать авторские права и приводить к юридическим последствиям. Важно определить, кто является владельцем прав на сгенерированное изображение — разработчик модели, алгоритм или кто-то другой.

Ещё одной серьёзной проблемой является возможность создания поддельных изображений и видео, известных как deepfake. Эти технологии могут использоваться для создания фальшивых новостей, компрометирующих видеозаписей и другой дезинформации, что может нанести значительный вред как отдельным людям, так и обществу в целом. Deepfake могут подрывать доверие к медиа, усложнять различение правды и лжи и использоваться в политических целях для манипуляции общественным мнением. Это требует разработки надежных методов для обнаружения поддельных материалов и создания законодательных рамок, регулирующих их использование.

Кроме того, нейросети для генерации изображений могут оказывать влияние на рынок труда. Автоматизация процессов создания контента может привести к сокращению рабочих мест в креативных индустриях, таких как дизайн, фотография и маркетинг. Это вызывает беспокойство среди профессионалов, которые могут оказаться замененными алгоритмами. С другой стороны, эти технологии также могут создавать новые рабочие места и возможности, требующие навыков работы с нейросетями и глубоким обучением [8].

Также стоит отметить проблему biases, или предвзятости, в данных, на которых обучаются нейросети. Если обучающие наборы данных содержат предвзятые или стереотипные изображения, модели могут воспроизводить эти предубеждения, что может усилить социальные неравенства и дискриминацию. Поэтому важно разрабатывать методы для выявления и устранения предвзятости в данных и алгоритмах, чтобы генеративные нейросети способствовали созданию более справедливого и инклюзивного общества [9, 10].

Этические и социальные аспекты использования нейросетей для генерации изображений требуют внимательного рассмотрения и постоянного контроля. Необходимы усилия по разработке нормативных и правовых рамок, которые будут регулировать использование этих технологий, защищая права и интересы всех участников процесса. Важно также продолжать общественные дискуссии и исследования, направленные на понимание и смягчение потенциальных негативных последствий, чтобы обеспечить ответственное и этичное использование генеративных нейросетей. Таким образом, эти технологии смогут приносить пользу обществу, оставаясь при этом безопасными и справедливыми.

Выводы

Нейросети для генерации изображений представляют собой одну из самых захватывающих и перспективных технологий в области искусственного интеллекта. Благодаря современным достижениям в области глубокого обучения, такие модели способны создавать фотореалистичные изображения, стилизованные произведения искусства, анимации и даже видео, открывая новые возможности для креативных индустрий и бизнеса.

В этой статье мы рассмотрели основные аспекты работы нейросетей, включая их обучение и архитектуры, такие как GAN, VAE и Diffusion Models. Мы также изучили различные примеры применения генеративных нейросетей в таких областях, как искусство, дизайн, медицина, развлечения и маркетинг. Эти технологии не только облегчают процесс создания контента, но и позволяют генерировать новые, уникальные решения, которые ранее были недоступны.

Однако, вместе с впечатляющими возможностями генеративных нейросетей возникают и значительные этические и социальные вопросы. Проблемы авторских прав, возможность создания поддельных изображений и видео (deepfakes), а также потенциальное воздействие на

рынок труда требуют внимательного рассмотрения и разработки соответствующих нормативных и правовых рамок.

Итак, нейросети для генерации изображений имеют огромный потенциал для трансформации различных отраслей, но важно подходить к их использованию с осознанием всех этических и социальных аспектов. Важно продолжать исследования и разработки в этой области, стремясь к созданию технологий, которые будут приносить пользу обществу, оставаясь при этом этичными и справедливыми.

Список литературы

1. Мария Жарова Как работает нейронная сеть: разбираемся с основами, 2024 // URL: <https://blog.skillfactory.ru/kak-rabotaet-nejronnaya-set-razbiraemsgya-s-osnovami/> (Дата обращения: 19.05.2024)
2. Нейросеть онлайн: погружение в мир искусственного интеллекта и нейросетей URL: <https://neuroseti-online.ru/> (Дата обращения: 20.05.2024)
3. Генеративно-сопоставительные связи Иэн Гудфеллоу, Жан Пуже-Абади, Мехди Мирза, Бинг Сюй, Дэвид Вард-Фарли, Шерджилозаир, Аарон Курвиль, Йошуа Бенгио Generative Adversarial Nets, 2014 // URL: https://papers.nips.cc/paper_files/paper/2014/hash/5ca3e9b122f61f8f06494c97b1afccf3-Abstract.html (Дата обращения: 20.05.2024)
4. Дидерик Пи Кингма, Макс Веллинг Auto-Encoding Variational Bayes, 2013// URL: <https://arxiv.org/abs/1312.6114> (Дата обращения: 22.05.2024)
5. Вероятностные модели диффузии с шумоподавлением Джонатан Хо, Аджай Джейн, Питер Эббил Denoising Diffusion Probabilistic Models, 2020 // URL: <https://arxiv.org/abs/2006.11239> (Дата обращения: 22.05.2024)
6. 8 лучших художественных генераторов искусственного интеллекта на базе GAN Перси Л. 8 Top AI Art Generators Powered by GAN, 2024 // URL: <https://www.ipic.ai/blogs/8-top-ai-art-generators-powered-by-gan/> (Дата обращения: 23.05.2024)
7. Список приложений генеративно-сопоставительных сетей Simplilearn List Of Generative Adversarial Networks Applications, 2023 // URL: <https://www.simplilearn.com/generative-adversarial-networks-applications-article> (Дата обращения: 23.05.2024)
8. Социально-этические проблемы, связанные с генеративным ИИ, 2023 Атуза Касирзаде // URL: <https://rse.org.uk/resources/resource/blog/the-socio-ethical-challenges-of-generative-ai/> (Дата обращения: 23.05.2024)
9. Булгакова А.В., Сафонова Т.В., Кутикова В.С. Классификация нейронных сетей Информационные технологии и системы: управление, экономика, транспорт, право. 2023. № 1 (45). С. 11-18.
10. Булгакова А.В., Сафонова Т.В., Диденко А.Ю. Этапы разработки и внедрения нейронной сети в проект Информационные технологии и системы: управление, экономика, транспорт, право. 2023. № 1 (45). С. 87-92.

References

1. Maria Zharova How a neural network works: understanding the basics, 2024 // URL: <https://blog.skillfactory.ru/kak-rabotaet-nejronnaya-set-razbiraemsgya-s-osnovami/> / (Date of access: 05/19/2024)

2. Neural network online: immersion in the world of artificial intelligence and neural networks
URL: <https://neyroseti-online.ru/> / (Date of request: 05/20/2024)
 3. Generative-adversarial relations Ian Goodfellow, Jean Pouget-Abadi, Mehdi Mirza, Bing Xu,
David Ward-Farley, Shergilozair, Aaron Courville, Joshua Bengio Generative Adversarial
Nets, 2014 // URL: https://papers.nips.cc/paper_files/paper/2014/hash/5ca3e9b122f61f8f06494c97b1afccf3-Abstract.html (Date of application: 05/20/2024)
 4. Diederik Pi Kingma, Max Welling Auto-Encoding Variational Bayes, 2013// URL: <https://arxiv.org/abs/1312.6114> (Date of application: 05/22/2024)
 5. Probabilistic noise-canceling diffusion Models Jonathan Ho, Ajay Jane, Peter Abill Denoising
Diffusion Probabilistic Models, 2020 // URL: <https://arxiv.org/abs/2006.11239> (Date of
application: 05/22/2024)
 6. 8 Best Artificial Intelligence Art Generators based on GAN Percy L. 8 Top AI Art Generators
Powered by GAN, 2024 // URL: <https://www.ipic.ai/blogs/8-top-ai-art-generators-powered-by-gan/> / (Date of request: 05/23/2024)
 7. List Of applications Of Generative Adversarial Networks Simplilearn List Of Generative
Adversarial Networks Applications, 2023 // URL: <https://www.simplilearn.com/generative-adversarial-networks-applications-article> (Date of application: 05/23/2024)
 8. Socio-ethical problems related to generative AI, 2023 Atuza Kasirzade // URL: <https://rse.org.uk/resources/resource/blog/the-socio-ethical-challenges-of-generative-ai/> / (Date
of request: 05/23/2024)
 9. Bulgakova A.V., Safonova T.V., Kutikova V.S. Classification of neural networks Information
technologies and systems: management, economics, transport, law. 2023. No. 1 (45). pp. 11-
18.
 11. Bulgakova A.V., Safonova T.V., Didenko A.Yu. Stages of development and implementation
of a neural network in the Information technologies and systems project: management,
economics, transport, law. 2023. No. 1 (45). pp. 87-92.
-