

ГЕНЕРАТИВНЫЕ НЕЙРОСЕТИ КАК ИНСТРУМЕНТ МОУШН-ДИЗАЙНЕРА

В статье рассматривается применение популярных генеративных нейросетей в моушн-дизайне и оценивается ряд их преимуществ и ограничений. Для работы с моушн-графикой основной упор сделан на нейросеть Stable Diffusion. Приводятся плюсы и минусы использования Stable Diffusion. Анализ нейросетей позволит лучше понять возможности использования генеративных инструментов и привлечь к ним больше внимания, определить перспективы разработки, доработки и доступности.

Ключевые слова: нейронные сети, моушн-дизайн, Stable Diffusion

D. D. Dudareva, A. A. Kozhedub

GENERATIVE NEURAL NETWORKS AS A MOTION DESIGNER'S TOOL

The article examines the use of popular generative neural networks in motion design and evaluates a number of their advantages and limitations. As for motion graphics, the main focus is on the Stable Diffusion neural network. The advantages and disadvantages of using the Stable Diffusion are presented. Analysis of neural networks will allow us to better understand the possibilities of using generative tools and draw more attention to them and determine the prospects for development, improvement and accessibility.

Keywords: neural networks, motion design, Stable Diffusion

Еще 10–15 лет назад статичное изображение было способно привлечь внимание. Однако сегодня такую графику успешно используют в качестве фона для видео и анимации. Моушн-дизайн оживляет графику более продвинутыми способами. Это анимация, помогающая дизайнеру сделать визуализацию интереснее, объяснить авторскую задумку. Моушн используют в сфере медиа-развлечений, видеоиграх, фильмах, веб-дизайне и т. д. Он показывает идею, не рассказывая полноценную историю, в отличие от анимации, которая содержит полноценный сюжет.

Приемы анимирования визуальных эффектов активно используют в киноиндустрии, рекламе и видеоклипах. В брендинге с помощью визуальных эффектов можно сделать подачу текста более интересной для пользователей. Активно динамическая графика включена в работу интерфейсов, сферу ui/ux и web дизайна.

Нейросеть построена по принципу организации человеческого мозга, именно поэтому их можно обучить. Нейронные сети умеют распознавать голос, текст, предметы, лица, водить машину, поддерживать разговор, рисовать картины, улучшать качество видео, анализировать действия на сайте, подбирать контент и т. д.

Выход эффективной нейросети определенного формата влечет за собой выход сходных инструментов. Например, после выхода ChatGPT появился поисковик Bink от Microsoft, который не только пишет достаточно качественные тексты, но и умеет искать картинки и видео в интернете. А выход Dalle2 приблизил встречу с новыми генеративными инструментами, которые рисуют картинки по текстовому запросу (Midjourney, Stable Diffusion, Dream).

Нейросети создают вирусные видео за считанные минуты при помощи нескольких промтов, также быстро они создают гиф-анимацию. Такой взрывной рост разнообразных технологий не может не сказаться на индустрии. Новые сервисы, вроде Vidyo, способны облегчить работу видеомонтажера. Сервисы Synthetik, Rokoko и WarpFusion эффективно

работают с моушн-дизайном и анимацией. Сегодня совсем не обязательно иметь отношение к киноиндустрии, ведь можно создавать видеоролики вместе с озвучкой, музыкой и монтажом, не выходя из дома. Также есть нейросети, которые улучшают качество видео, делают мастеринг. Очень хорошо искусственный интеллект помогает в поиске музыки, сокращая время на долгий подготовительный процесс. Нейросеть может нарезать какие-то интересные кадры, а моушн-дизайнер будет оценивать результат и собирать окончательный вариант. Нейросети способны быстро объединить видеофрагменты, определить лучшие кадры, удалить дубликаты или неудачные дубли. Нейросеть — это просто вспомогательный инструмент, она не может обеспечить полный цикл создания продукта.

Как же нейросети работают? Представим, что у нас есть вилка и ложка, нам надо научить компьютер отличать их. Сначала мы показываем ему много примеров вилок и ложек и говорим, что это такое. Нейросеть обрабатывает эти примеры и настраивается таким образом, чтобы лучше различать отличия между ними. Чем больше изображений мы разбираем таким способом, тем лучше нейросеть будет различать данные предметы. Нейросеть разбирает изображение на пиксели, прогоняет через свои нейроны или синапсы и выдает результат. Нейросеть отличается от искусственного интеллекта. Это всего лишь один из методов обработки информации. Но для удобства искусственный интеллект, который использует данный метод, называют нейросетью. Явными преимуществами нейросетей являются сбор и анализ данных, обработка больших объемов информации быстрее человека. Это делает ее идеальным инструментом для любого задания, связанного с анализом данных и машинным обучением. Нейросеть автоматизирует рутинные задачи, находит скрытые закономерности в данных, которые могут быть незаметны для человека и на основе этого делает прогнозы и рекомендации.

Некоторые нейросети пользуются особой популярностью. Midjourney рисует красивые арты. Он выдает наиболее кинематографичные, красивые, детальные работы, например, для создания мудбордов, получения референсов. Но в работе с ней есть свои минусы: можно работать только через дискорд, мало параметров, которые можно настраивать, она не умеет работать с видео.

Chat GPT используется для генерации сценариев, когда необходимо придумать тему и текст для влога, написать скрипт для эффекта. Это мощный инструмент генерации текста, кода или алгоритма. С ним можно общаться как с реальным человеком.

Descript позволяет монтировать пользователю, не имеющему даже базовых навыков монтажа. Программа анализирует исходник и удаляет все лишнее, монтируя готовое видео.

Nerf от Nvidia — это нейросеть, которую обучили смотреть так, как смотрят люди, деконструировать объекты и окружающее пространство на основе всего нескольких фото.

Stable Diffusion — это популярная бесплатная нейросеть с открытым кодом, которая позволяет генерировать картинки на основе запроса. Программа производительна за счет вывода изображений высокого разрешения, быстрой скорости генерирования, низких требований к ресурсам памяти. Работы, сгенерированные Stable Diffusion, могут быть реалистичными, абстрактными или вообще выполненными в стиле определенного художника.

Благодаря открытому коду вокруг Stable Diffusion формируется огромное активное комьюнити, появляются приложения по генерации картинок онлайн, на телефоне, через облако или локально, если есть хорошая видеокарта. Теперь даже не обязательно придумывать свой текстовый запрос. В сети можно подобрать подходящий промпт или сид.

Функции нейросетей:

- генерация картинок по текстовому запросу;
- редактирование фото;
- генерация видео;
- генерация фотореалистичного изображения с использованием собственного рисунка.

Для моушн-дизайна с помощью Stable Diffusion можно сгенерировать изображение по своему запросу и применить его где угодно. Создание картинок по текстовому запросу в

разы ускорит работу дизайнера. Но бесплатная версия, к сожалению, имеет весьма ограниченный функционал. В бесплатной версии нужно задавать негативный запрос, то есть вводить предметы, которых не должно быть на работе.

Существует еще и десктопная версия. В ней уже можно выбрать любой размер картинки, также выбор инструментов гораздо больше.

Stable Diffusion генерирует по алгоритму текст-картинка, картинка-картинка, текст-видео, картинка-видео. Доступна функция гибкого изменения изображения.

Чтобы получить желаемый запрос от Stable Diffusion, нужно правильно задать вопрос и поговорить с нейросетью на ее языке — это называется текстовый промт. Это текстовая подсказка для нейросети, написанная на естественном языке, и наш запрос к ней. Каждый добавленный промт сужает спектр поиска и конкретизирует результат. Получается, чтобы получить изображение, приближенное к нашей задумке, надо указывать самые очевидные и неочевидные промты, вплоть до фокусного расстояния объектива, позы персонажа и др. У нейросетей есть огромный запас данных — dataset с предметами, сюжетами, стилями, разными материалами, из которых могут быть сделаны предметы. Когда мы делаем запрос к нейросети, она обращается к своему dataset и синтезирует для нас изображение. Структура промта:

- сюжетная (то, что нарисовано на картинке);
- изобразительная (то, в каком стиле или в каком материале сделано);
- техническая (качество изображения).

Существует целая наука — промт-инжиниринг. Однако даже после вывода готового результата может потребоваться корректировка изображения.

Многие думают, что результат, который выдают нейросети, предсказуем. На самом деле это не совсем так. Но по большому счету результат будет очевиден, так как человек задает конкретную поставленную задачу.

Получается, что нейронные сети могут быть инструментом для решения многих сложных задач, но требуют проверки для достижения поставленной задачи. Они не заменят человека, так как делают свои работы на основе уже существующих. Кроме того, минусы нейросетей в том, что они требуют больших объемов данных для обучения корректной работе, что может быть сложно для многих задач. Искусственный интеллект может быть ограничен в своей способности адаптироваться к новым задачам и изменениям окружающей среды. Несомненно, со временем этот вопрос будет решен, но для этого нужно, чтобы нейросеть успела изучить каждую сферу человеческой деятельности. А вот задачи, связанные с пониманием человеческих эмоций и управлением ими, являются практически невыполнимыми для нейросети.

Высказывания о том, что нейросети заменят человека, появились из-за недостаточной осведомленности о работе искусственного интеллекта. Нейронные сети не смогут заменить креативность и инновационность, которые являются ключевыми факторами в различных сферах, таких как искусство, музыка, дизайн и наука. Нейросети и алгоритмы могут создавать новые идеи и продукты, но они не могут заменить творческий процесс и воображение человека, которое позволяет создавать что-то уникальное и ранее не встречающееся. При этом невероятная способность нейросетей решать мелкие задачи просто феноменальна.

Пару лет назад Midjourney выдавал смешные и нелепые результаты, но время шло, нейросеть совершенствовалась и училась на своих ошибках. Теперь она создает шедевры. Качество картинок не отличишь от реалистичных изображений, сделанных известными художниками.

Если задачу генерации статичных изображений можно считать уже решенной, то ситуация с анимацией или видео немного сложнее. Одна из лучших нейросетей для создания анимации — Deforum Stable Diffusion. Это серьезный инструмент со множеством опций, которые позволяют настраивать количество кадров, движение камеры, зум, интерполяцию и другие. Нейросеть доступна через сервис Google Colab. Начать следует с генерации

статичных изображений. Набив руку на промтах, поняв нюансы стилей и композиции, реально превратить искусственный интеллект в свой рабочий инструмент. ИИ хорошо справляется с автоматическим монтажом. Требование к органичности и однородности объектов при их движении на видео приводит к тому, что объекты на каждом отдельном кадре становятся более правдоподобными. Это достигается путем обучения на коротких видео, что более эффективно, чем если бы сеть училась на статических изображениях. И если продолжать тренировать генеративные нейросети на качественных видео, то динамика скачка в более высокую производительность будет весьма существенной. Надо лишь подождать, когда ресурсы будут это позволять.

Также ожидаются в будущем модели для генерации 3D-объектов по текстовому запросу. На данный момент направление работы с 3D-объектами самое сложное, но уже есть такие нейросети как Diffusion.

Очень скоро Chat GPT сможет воспринимать не только текстовую информацию, но и графическую, то есть можно будет загрузить какое-нибудь изображение или видео и пообщаться с нейросетью на тему содержимого этого медиафайла.

Нейросети не заменят человека, они будут служить инструментами, помогающими решать задачи более эффективно и быстро. Человеческий фактор, эмоции и творческий процесс всегда будут играть важную роль, и в этом искусственный интеллект не сможет превзойти человека.

ЛИТЕРАТУРА

1. Аггарвал Чару А. Нейронные сети и глубокое обучение: учебный курс.: пер. с англ. СПб.: ООО «Диалектика», 2020. 752 с.
2. Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение: пер. с англ. А. А. Слинкина. 2-е изд., испр. М.: ДМК Пресс, 2018. 652 с.
3. Нейросети в генерации видео. URL: <https://habr.com/ru/companies/ruvds/articles/716660/> (дата обращения 03.10.2023).
4. Николенко С., Кадурин А., Архангельская Е. Глубокое обучение. СПб.: Питер, 2018. 480 с.
5. Решетникова М. Stable Diffusion: как нейросеть создает работы в стиле киберпанк и Disney: электронный ресурс (дата обращения 29.09.2023).
6. Фостер Д. Генеративное глубокое обучение. Творческий потенциал нейронных сетей. испр. СПб.: Питер, 2020. 336 с.
7. Хайкин С. Нейронные сети. Полный курс. 2-е изд.: пер. с англ. М.: ООО «И. Д. Вильямс», 2016. 1104 с.

Сведения об авторах:

Дударева Диана Дмитриевна, старший преподаватель, кафедра дизайна и моды, Витебский государственный технологический университет; dudareva_d@mail.ru

Кожедуб Анастасия Андреевна, студентка, 2 курс, Витебский государственный технологический университет; anastasiakozhedub1@gmail.com

About the authors:

Dudareva Diana Dmitrievna, Senior Lecturer, Department of Design and Fashion, Vitebsk State Technological University; dudareva_d@mail.ru

Kozhedub Anastasiya Andreevna, student, 2nd year, Vitebsk State Technological University; anastasiakozhedub1@gmail.com