

УДК 004.8

ПЕРЕДОВЫЕ ТЕХНОЛОГИИ В СОЗДАНИИ ВИЗУАЛЬНОГО КОНТЕНТА: ВЗГЛЯД ЧЕРЕЗ ПРИЗМУ ГЕНЕРАТИВНО-СОСТЯЗА- ТЕЛЬНЫХ НЕЙРОСЕТЕЙ

Гуркина А.О., студент гр. 010402, I курс
Научный руководитель: Кубанских О.В., к.ф.-м.н., доцент
Брянский государственный университет имени академика
И.Г. Петровского, г. Брянск
г. Брянск

В настоящее время нейронные сети играют ключевую роль в сфере обработки и создания изображений. Они обладают способностью не только анализировать и классифицировать визуальный контент, но и творчески преобразовывать его. Одним из наиболее значимых видов нейронных сетей в данном контексте являются генеративно-сопоставительные нейросети (GAN), с помощью которых можно создавать уникальный и качественный визуальный контент, что делает его весьма перспективным для различных областей исследований и применений, включая компьютерное зрение, искусственный интеллект и графический дизайн.

Структура GAN состоит из следующих элементов: генератор и дискриминатор, которые работают вместе в процессе обучения для создания новых данных, неотличимых от реальных. Рассмотрим структуру и процесс обучения генеративно-сопоставительных нейронных сетей, который представлен на рисунке 1, более подробно [1].

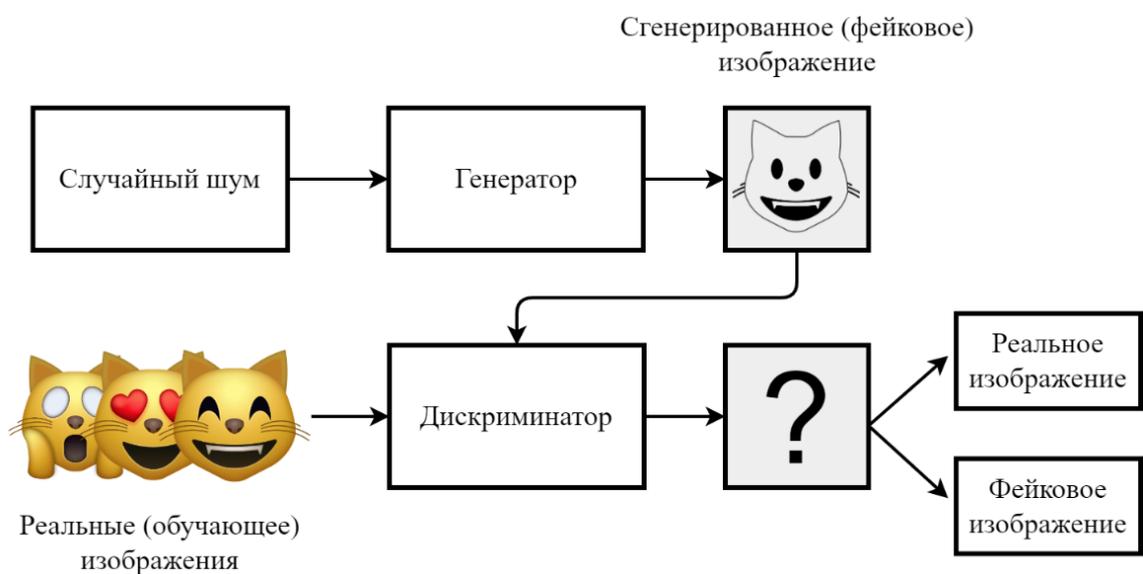


Рисунок 1 – Процесс обучения генеративно-сопоставительных нейросетей

Генератор в контексте машинного обучения представляет собой модель, которая принимает на вход случайный шум (или вектор) и генерирует данные, которые должны быть похожи на реальные. Цель генератора состоит в том, чтобы создавать данные, которые становятся все более похожими на реальные, с тем чтобы обмануть другую модель, называемую дискриминатором [5].

Дискриминатор представляет собой еще один значимый элемент системы GAN. Он принимает на вход данные (из обучающего набора) или сгенерированные генератором, и пытается определить их истинное происхождение. Задача дискриминатора – правильно идентифицировать поддельные данные [3].

В процессе обучения GAN используется функция потерь, которая играет ключевую роль в оптимизации работы обоих компонентов - генератора и дискриминатора (формула 1).

$$L = \frac{1}{N} \sum_i L_i(f(x_i, W), y_i) \quad (1)$$

где L – это средняя сумма потерь для всех обучающих данных, f – классификатор, который принимает на вход пример x_i , весовую матрицу W и делает прогноз для класса y_i . Функция потерь L_i выполняет сравнение между предсказанным классом и истинным классом для данного примера. Это сравнение позволяет определить, насколько точно или неточно модель классификатора классифицирует данный пример x_i [7].

Цель заключается в минимизации функции потерь L_i для улучшения производительности классификатора f . А именно, насколько успешно генератор «обманывает» дискриминатор, создавая данные, которые трудно отличить от реальных, и насколько успешно дискриминатор распознает поддельные данные.

По мере процесса обучения генератор становится все более и более «утонченным» в создании реалистичных данных, в то время как дискриминатор совершенствует свои навыки классификации, чтобы эффективно отличать поддельные данные от реальных. Этот взаимодействующий процесс обновления и улучшения обеих сетей способствует достижению равновесия и повышению качества новых генерируемых данных.

Обучение GAN продолжается до тех пор, пока генератор не достигнет уровня, когда созданные им данные становятся практически неотличимыми от реальных. Этот результат является индикатором успешного обучения GAN и обычно является конечной целью процесса обучения [4].

Таким образом, применение GAN открывает удивительные возможности для создания уникального и выразительного визуального контента. Этот тип нейронных сетей способен генерировать реалистичные изображения, не существующие в реальном мире, обладающие творческим потенциалом. Благодаря возможности обучения на больших объемах данных, GAN способны создавать удивительно реалистичные текстуры, изображения и даже видео.

Например, в задаче генерации фотореалистичных изображений людей нейронные модели способны создавать уникальные изображения лиц, которые выглядят так, будто они настоящие фотографии людей (рис.2).



Рисунок 2 – Результат работы нейросети по текстовому запросу «Обложка журнала с женщиной»

Аналогично, GAN могут использоваться для создания новых пейзажей или даже целых сцен. На рисунке 3 генератор создал изображение лица актера Райана Гослинга в локации из книги «Ведьмак», которое не существовало никогда ранее, и оно выглядит достаточно реалистично [2].

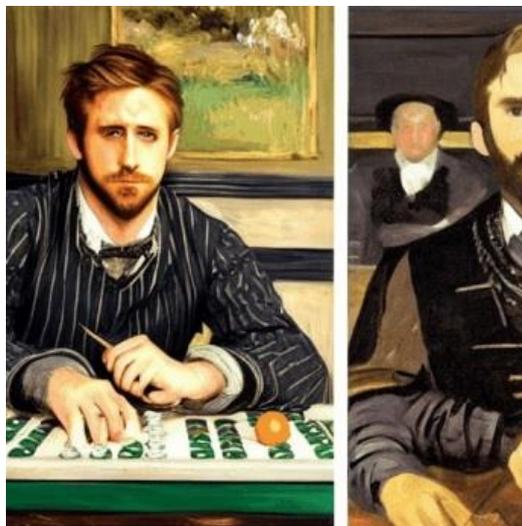


Рисунок 3 – Результат работы нейросети по текстовому запросу «Картина маслом в стиле Эдуарда Мане. Райан Гослинг в доспехах ведьмака из «Школы кошек» играет в нарды с Толкином в таверне из «Ведьмака»»

Генерация новых изображений с помощью GAN открывает возможности для создания контента, который ранее не существовал, и может быть полезным в различных областях, таких как разработка игр, визуальные эффекты, создание искусства и дизайн.

Нейросети генеративно-согласительного типа могут применяться для имитации стилей при обработке изображений. Например, реалистичное фото может быть стилизовано под аниме-мультфильм (рис. 4).

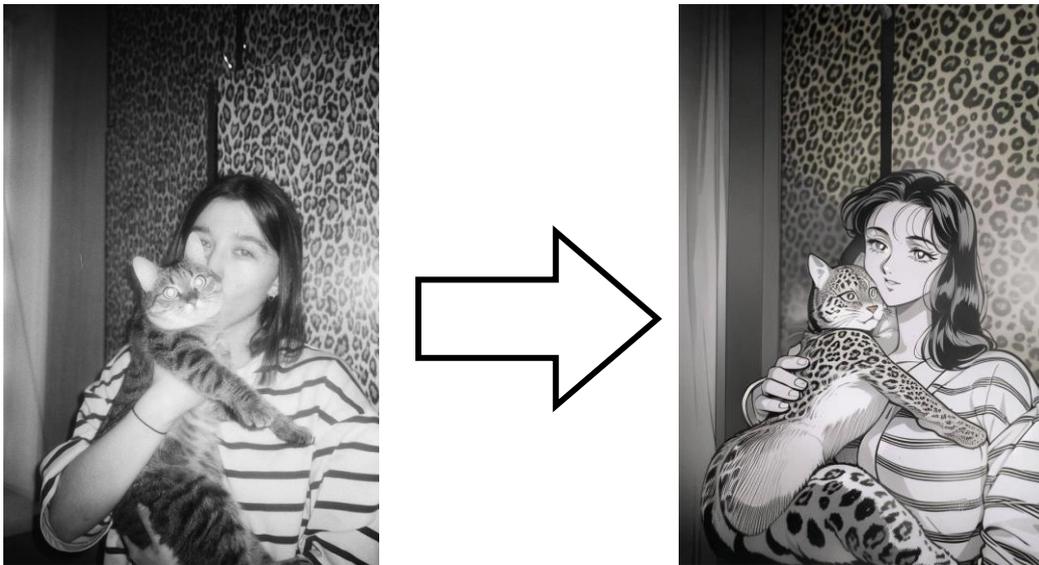


Рисунок 4 – Результат работы нейросети в стилизации реалистичной фотографии под аниме

Этот процесс называется «переносом стиля». С помощью GAN можно обучить модель, которая способна анализировать содержание и стиль двух разных изображений и затем сгенерировать новое изображение, сочетающее содержание одного изображения и стиль другого. Этот подход открывает возможности для творчества и экспериментов с изображениями [4].

Таким образом, использование генеративно-согласительных нейронных сетей в создании визуального контента дает новые возможности для художников, дизайнеров и разработчиков. Они позволяют автоматизировать процессы творчества, экспериментировать с новыми стилями и формами, а также создавать уникальные и впечатляющие визуальные произведения.

В заключение, важно подчеркнуть, что генеративно-согласительные нейросети представляют собой мощный инструмент в создании визуального контента, который не только может изменить представление о творчестве, но и открыть новые горизонты для развития искусства и дизайна.

Список литературы:

1. Гуркина, А. О. Генеративно-согласительные нейросети и их применение в области обработки изображений / А. О. Гуркина // Цифровые, компьютерные и информационные технологии в науке и образовании : Сборник статей Межрегиональной научно-практической конференции с международным участием, Брянск, 01–02 ноября 2023 года. – Брянск: Брянский государственный университет имени академика И.Г. Петровского, 2023. – С. 331-335.
2. Гуркина, А. О. Использование циклически согласованных согласительных сетей для преобразования изображения / А. О. Гуркина, О. В. Кубанских

// Россия молодая : СБОРНИК МАТЕРИАЛОВ XIV ВСЕРОССИЙСКОЙ, НАУЧНО-ПРАКТИЧЕСКОЙ КОНФЕРЕНЦИИ МОЛОДЫХ УЧЕНЫХ С МЕЖДУНАРОДНЫМ УЧАСТИЕМ, Кемерово, 18–21 апреля 2023 года. – Кемерово: Кузбасский государственный технический университет имени Т.Ф. Горбачева, 2023. – С. 31717.1-31717.5. – EDN JIGAPN.

3. Гуркина А.О. История развития GAN-сетей. Перспективы использования их в бизнес-сфере // Материалы XVI Международной студенческой научной конференции «Студенческий научный форум» URL: <https://scienceforum.ru/2024/article/2018036173>

<https://scienceforum.ru/2024/article/2018036173> (дата обращения: 21.03.2024).

4. Муаль Мутуама Нда Бьенвеню ИСПОЛЬЗОВАНИЕ ПРЕДОБУЧЕННОЙ НЕЙРОСЕТИ (VGG16) ДЛЯ РЕШЕНИЯ ЗАДАЧИ ПЕРЕНОСА СТИЛЯ ИЗОБРАЖЕНИЯ // Современные информационные технологии и ИТ-образование. 2022. №2. URL: <https://cyberleninka.ru/article/n/ispolzovanie-predobuchennoy-neyroseti-vgg16-dlya-resheniyazadachi-perenosy-stilya-izobrazheniya> (дата обращения: 22.03.2024).

5. Намиот Д.Е., Ильюшин Е.А. ПОРОЖДАЮЩИЕ МОДЕЛИ В МАШИННОМ ОБУЧЕНИИ // International Journal of Open Information Technologies. 2022. №7. URL: <https://cyberleninka.ru/article/n/porozhdayuschie-modeli-v-mashinnom-obuchenii> (дата обращения: 25.03.2024).

6. Стэнфордский курс: лекция 3. Функция потерь и оптимизация: сайт. – URL: <https://www.reg.ru/blog/stehnfordskij-kurs-lekciya-3-funkciya-poter-i-optimizaciya/> (дата обращения: 23.03.2024). – Текст: электронный.