

Использование мультимодальной нейросетевой платформы Perplexity для генерации фотореалистичных изображений и видеоряда на основе текстовых промтов

Анишкова Анастасия Сергеевна

Приамурский государственный университет имени Шолом-Алейхема

Студент

Аннотация

Целью исследования является разработка и практическая апробация процесса генерации фотореалистичных изображений и видеоряда с использованием мультимодальной нейросетевой платформы Perplexity, основанного на применении текстовых промтов для точного управления стилистическими, композиционными и эмоциональными характеристиками визуального контента. Выбор нейросетевой платформы Perplexity обусловлен её возможностями по синтезу фотореалистичных изображений на основе сочетания текстовых описаний (промтов) и визуального референса (исходной фотографии). Применение такого гибридного подхода обеспечивает, с одной стороны, высокую точность воспроизведения индивидуальных черт модели, а с другой - позволяет гибко управлять художественными параметрами композиции, освещения и стиля в соответствии с заданными творческими задачами.

Ключевые слова: изображение, промт, нейросеть, Perplexity.

Using the Perplexity multimodal neural network platform to generate photorealistic images and videos based on text prompts

Anishkova Anastasia Sergeevna

Sholom-Aleichem Priamurskiy State University

Student

Abstract

Research Objective The research objective is to develop and practically test the process of generating photorealistic images and video sequences using the Perplexity multimodal neural network platform, which is based on the use of text prompts to accurately control the stylistic, compositional, and emotional characteristics of visual content. The choice of the Perplexity neural network platform is based on its ability to synthesize photorealistic images using a combination of text descriptions (prompts) and visual references (original photographs). This hybrid approach ensures high accuracy in reproducing the individual features of the model, while also allowing for flexible control over the artistic parameters of composition, lighting, and style, depending on the specific creative objectives.

Keywords: image, prompt, neural network, Perplexity.

1 Введение

1.1 Актуальность

Актуальность темы определяется активным развитием нейросетевых технологий и их интеграцией в процессы создания визуального контента. Применение мультимодальных платформ, таких как Perplexity, позволяет формировать новые подходы к генерации изображений и видео, основанные на взаимодействии текстовых описаний и визуальных референсов, что особенно важно для современного цифрового медиапространства.

1.2 Обзор исследований

Е.В. Ильинская, Е.Н. Голышева, А.А. Медведев, Н.С. Масалитин описали применение генеративно-сопоставительных нейросетей для генерации изображений [1]. Способы генерации изображений при помощи искусственных нейронных сетей применили И. С. Шустов, Ю. С. Белов [2]. Е. В. Кагарлык, Ю. Ю. Балахонов создали картинки и описали обработку в новых технологиях нейронной сети [3]. Генеративно-сопоставительные сети для генерации изображений по тексту использовали А. О. Левин, Ю. С. Белов [4]. И. С. Шустов, Ю. С. Белов подготовили и обработали данные в задачах генерации фотореалистичных изображений [5].

1.3 Цель исследования

Целью исследования является разработка и практическая апробация процесса генерации фотореалистичных изображений и видеоряда с использованием мультимодальной нейросетевой платформы Perplexity, основанного на применении текстовых промтов для точного управления стилистическими, композиционными и эмоциональными характеристиками визуального контента.

2 Материалы и методы

В качестве инструментария исследования была использована мультимодальная нейросетевая платформа Perplexity, основанная на технологии генерации изображений, совмещающей визуальный референс (исходную фотографию) и текстовый запрос (промпт) для точного управления параметрами и характеристиками итогового результата.

3 Результаты

Генерация изображений осуществлялась с использованием мультимодальной нейросетевой платформы Perplexity. На первом этапе в систему загружалось референсное изображение (рис. 1), которое использовалось в качестве основы для определения стилистических и композиционных параметров итоговых визуальных решений.



Рисунок 1. Референс

С помощью детально разработанного промта было получено изображение, представляющее собой ультрареалистичный черно-белый студийный портрет женщины, созданный в эстетике Питера Линдберга — с мягким направленным освещением, естественной фактурой кожи и выразительной, подлинно человеческой атмосферой кадра (рис.2-5).

Промт: Ultra-realistic black-and-white studio portrait photograph of a confident woman leaning gracefully on the backrest of a modern minimal chair. She wears a sleek, textured black outfit — fine knit or tailored wool — with subtle fabric detail. Captured from a low-angle perspective to accentuate her strong facial structure, defined cheekbones, and intense, self-assured gaze. Her hair is natural, softly flowing, with individual strands visible. Lighting is cinematic: soft directional key light from upper left, gentle backlight for separation. Background is pure matte black. Image is rendered in extreme high resolution (8K), with hyper-detailed textures in skin, eyes, hair, and clothing. Photographic style inspired by Peter Lindbergh — timeless, raw, elegant, with natural skin tones preserved in grayscale. No makeup, no retouching, authentic human presence (Ультрареалистичная черно-белая студийная портретная фотография

уверенной в себе женщины, грациозно откинувшейся на спинку современного стула в стиле минимализм. На ней элегантный черный костюм из тонкой трикотажной ткани или шерсти, сшитой на заказ, с изящными деталями из ткани. Снимок сделан с низкого ракурса, чтобы подчеркнуть выразительность ее лица, четко очерченные скулы и пристальный, уверенный взгляд. Ее волосы натуральные, мягко струящиеся, с видимыми отдельными прядями. Освещение кинематографическое: мягкий направленный свет в верхнем левом углу, мягкая подсветка для разделения. Фон - чистый матово-черный. Изображение выполнено в чрезвычайно высоком разрешении (8K), с очень детализированными текстурами кожи, глаз, волос и одежды. Фотографический стиль, вдохновленный Питером Линдбергом, — неподвластный времени, лаконичный, элегантный, с естественными оттенками кожи, сохраненными в оттенках серого. Никакого макияжа, никакой ретуши, подлинное человеческое присутствие).



Рисунок 2. Первый вариант



Рисунок 3. Второй вариант



Рисунок 4. Третий вариант



Рисунок 5. Четвертый вариант

Для последующей генерации видеоряда были использованы ранее полученные фотографии. Процесс осуществлялся путём применения нескольких текстовых промтов, каждый из которых задавал уникальные визуальные и эмоциональные параметры сцены.

Первый промт описывал черно-белое студийное портретное видео с участием молодой женщины, сидящей на чёрном стуле в затемнённой студии, к которой подходит Санта-Клаус и аккуратно надевает на неё пушистую рождественскую шапку. Визуальный образ строился на сочетании мягкого направленного освещения, высокой контрастности и текстуры киноплёнки, создающих атмосферу уюта и праздничного настроения. Ссылка видео https://disk.yandex.ru/i/NNUQ3Bw_hH7HGg.

Промт: Black and white cinematic portrait video of a beautiful young woman with long straight hair sitting on a black chair in a dark studio. Behind her, Santa Claus gently places a fluffy white Christmas hat on her head. She starts with a neutral expression, then smiles broadly, laughs joyfully, and looks to the side. Soft dramatic lighting, high contrast, shallow depth of field, film grain texture. Static camera, medium close-up shot. Festive, warm, emotional atmosphere. color, cartoon, anime,

low quality, blurry, deformed hands, extra limbs, text, watermark, logo, bright background, overexposed (Черно-белое кинематографическое видео с портретом красивой молодой женщины с длинными прямыми волосами, сидящей на черном стуле в темной студии. Позади нее Санта-Клаус осторожно надевает ей на голову пушистую белую рождественскую шапочку. Она начинает с нейтрального выражения лица, затем широко улыбается, радостно смеется и отводит взгляд в сторону. Мягкое драматическое освещение, высокая контрастность, небольшая глубина резкости, зернистость пленки. Статичная камера, средний крупный план. Праздничная, теплая, эмоциональная атмосфера. цветной, мультфильм, аниме, низкое качество, размытое изображение, деформированные руки, лишние конечности, текст, водяной знак, логотип, яркий фон, передежженный)

Второй промт развивал концепцию, включая уточнённые характеристики освещения, глубины резкости и мимической выразительности модели. Итоговое изображение отличается кинематографичностью, реалистичной передачей эмоций и эстетикой, близкой к классическим черно-белым фильмам. Ссылка на видео <https://disk.yandex.ru/i/2N1ET4QZRBhLCQ>.

Промт: Cinematic black-and-white portrait video of a beautiful young woman with long, straight, silky hair sitting gracefully on a sleek black modern chair in a dimly lit studio. Behind her, Santa Claus — dressed in classic red velvet suit with white fur trim — gently and lovingly places a fluffy white Christmas hat on her head. She begins with a calm, neutral expression, then slowly breaks into a wide, radiant smile, bursts into joyful laughter, and turns her gaze softly to the side — as if reacting to someone off-camera. Soft, dramatic studio lighting: key light from front-left creates gentle highlights on her face and hair, rim light separates her from the deep black background. Shallow depth of field blurs the background subtly, keeping focus on her emotional expression. Film grain texture adds nostalgic, vintage cinematic feel. Static camera, medium close-up shot (chest up), 16:9 aspect ratio. Atmosphere: warm, festive, intimate, emotionally rich — like a holiday film still come to life. No color, no cartoon, no anime style — photorealistic, high detail, natural human proportions. Avoid: deformed hands, extra limbs, blurry faces, overexposure, text, logos, watermarks, bright backgrounds (Кинематографическое черно-белое портретное видео красивой молодой женщины с длинными, прямыми, шелковистыми волосами, грациозно сидящей на изящном черном современном стуле в тускло освещенной студии. Позади нее Санта—Клаус, одетый в классический красный бархатный костюм с белой меховой оторочкой, нежно и любовно надевает ей на голову пушистую белую рождественскую шапочку. Она начинает со спокойного, нейтрального выражения лица, затем медленно расплывается в широкой, лучезарной улыбке, заливается радостным смехом и мягко отводит взгляд в сторону — как будто реагирует на кого-то за кадром. Мягкое, эффектное студийное освещение: основной свет спереди слева создает нежные блики на ее лице и волосах, а свет в обрамлении выделяет ее из темно-черного фона. Небольшая глубина резкости слегка размывает фон, позволяя сфокусироваться на выражении ее эмоций. Зернистость пленки придает фильму ностальгический,

винтажный кинематографический оттенок. Статичная камера, крупный план среднего размера (по грудь), соотношение сторон 16:9. Атмосфера: теплая, праздничная, интимная, эмоционально насыщенная — как будто праздничный фильм все еще оживает).

4 Выводы

В ходе проделанной работы был реализован процесс генерации визуальных материалов с использованием мультимодальной нейросетевой платформы Perplexity. На основе загруженного референсного изображения и тщательно сформулированных текстовых промтов были получены ультрареалистичные черно-белые портреты, отличающиеся высокой детализацией и художественной выразительностью.

Дальнейшее применение этих изображений в качестве исходных данных позволило осуществить генерацию видеоряда, в котором удалось сохранить стилистическую целостность, эмоциональную атмосферу и характер освещения, заданные на этапе формирования исходных промтов.

Полученные результаты демонстрируют потенциал современных нейросетевых технологий для создания фотореалистичных изображений и видео в художественно-студийной эстетике, а также подтверждают эффективность использования текстовых описаний (prompt-based generation) как инструмента точного управления визуальным стилем и композицией в процессе медиапроизводства.

Библиографический список

1. Ильинская Е.В., Голышева Е.Н., Медведев А.А., Масалитин Н.С. Применение генеративно-сопоставительных нейросетей для генерации изображений // Научный результат. Информационные технологии. 2024. №1. С. 73-78.
2. Шустов И.С., Белов Ю.С. Способы генерации изображений при помощи искусственных нейронных сетей // Научно-технические технологии в приборостроении и развитии инновационной деятельности в вузе. 2016. С. 198.
3. Кагарлык Е. В., Балахонов Ю. Ю. Создание картинки и обработка её в новых технологиях нейронной сети //Современные информационные технологии в образовании, науке и промышленности. Искусственный интеллект в создании картин. 2021. С. 198-201.
4. Левин А. О., Белов Ю. С. Использование генеративно-сопоставительных сетей для генерации изображений по тексту // Научное обозрение - технические науки. 2023. №2. С. 10.
5. Шустов И. С., Белов Ю. С. Подготовка и обработка данных в задачах генерации фотореалистичных изображений //Международный студенческий научный вестник. 2018. №. 5. С. 139-139.