

Мультимодальный подход в генерации аватаров

Анишкова Анастасия Сергеевна

Приамурский государственный университет имени Шолом-Алейхема

Студент

Аннотация

Цель исследования — проанализировать влияние структурированных текстовых промтов на генерацию стилистически разнообразных портретных изображений мультимодальной нейросетью Perplexity при использовании единого референсного изображения. Для реализации выбрана нейросетевая платформа Perplexity, обеспечивающая генерацию фотореалистичных изображений на основе текстовых запросов и объединения фотографий. Такой подход обеспечивает создание высококачественных, реалистичных изображений, максимально соответствующих заданным параметрам, а также эффективное использование исходных фотографий для достижения точности и достоверности визуального результата.

Ключевые слова: изображение, промпт, нейросеть, Perplexity.

A multimodal approach to avatar generation

Anishkova Anastasia Sergeevna

Sholom-Aleichem Priamurskiy State University

Student

Abstract

The purpose of the study is to explore and analyze the impact of structured text prompts on the generation of stylistically diverse portrait images by the Perplexity multimodal neural network using a single reference image. The Perplexity neural network platform has been chosen for implementation, as it enables the generation of photorealistic images based on text queries and the combination of photographs. This approach ensures the creation of high-quality, realistic images that closely match the specified parameters, while effectively utilizing the original photographs to achieve accurate and reliable visual results.

Key words: image, prompt, neural network, Perplexity.

1 Введение

1.1 Актуальность

Актуальность исследования обусловлена стремительным развитием и доступностью генеративных нейросетей, которые в 2025 году стали ключевым инструментом для создания и управления цифровой идентичностью в социальных сетях и профессиональной среде. В этих условиях возрастает потребность в понимании механизмов управления генерацией изображений.

Ваше исследование напрямую отвечает на этот запрос, демонстрируя, как с помощью структурированных промтов можно целенаправленно влиять на художественный стиль и конечный результат, что является практически значимым навыком для пользователей и специалистов.

1.2 Обзор исследований

Е.В. Ильинская, Е.Н. Гольшева, 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

Цель исследования — проанализировать влияние структурированных текстовых промтов на генерацию стилистически разнообразных портретных изображений мультимодальной нейросетью Perplexity при использовании единого референсного изображения.

2 Материалы и методы

В качестве инструментария для данного исследования была выбрана мультимодальная нейросетевая платформа Perplexity. Использовалась технология генерации изображений, сочетающая визуальный референс (исходное фото) и текстовый запрос (промт) для управления конечным результатом.

3 Результаты

Генерация аватарки для социальных сетей производилась с использованием мультимодальной нейросетевой платформы Perplexity. Процесс начинался с загрузки в систему референсного изображения (рис. 1), которое служило основой для определения стилистических и композиционных характеристик конечных результатов.



Рисунок 1. Референс

Первый промпт был направлен на создание портрета в стиле неон-нуар (рис.2). Запрос описывал молодую женщину в черном кожаном плаще на фоне темного стекла с отражениями неоновых вывесок. Целью было добиться кинематографичной, таинственной и стильной атмосферы, напоминающей кадр из триллера. Особое внимание уделялось контровому свету, который должен был акцентировать лицо и волосы модели, создавая драматичную игру теней.

Промпт: Create an image of me. The portrait is in a cinematic neon noir style: a young woman with long blonde hair is half-turned to the camera, looking over her shoulder with a confident, focused expression. She is wearing a black leather coat, which highlights the structure and highlights of the material in the warm backlight. The background is dark glass with reflections and out-of-focus neon signs in red and orange tones, creating a tense, atmospheric depth to the frame. The light accentuates the model's face and hair, creating a dramatic play of shadows and reflections. The atmosphere is mysterious, stylish, and cinematic, like a scene from a trendy thriller or an urban-themed TV series. The photo should be taken with a modern camera

equipped with a 35mm lens and 4KHD quality (Создай изображение меня. Портрет выполнен в кинематографичном стиле неон-нуара: молодая женщина с длинными светлыми волосами стоит в полоборота к камере, глядя через плечо с уверенным, сосредоточенным выражением. Она одета в чёрный кожаный плащ, подчёркивающий структуру и блики материала в контровом тёплом свете. Фон — тёмное стекло с отражениями и расфокусированными неоновыми вывесками в красных и оранжевых тонах, создающими напряжённую, атмосферную глубину кадра. Свет акцентирует лицо и волосы модели, создавая драматичную игру теней и отблесков. Атмосфера — таинственная, стильная и кинематографичная, как кадр из модного триллера или авторского сериала в урбанистической эстетике. Фото должно быть сделано на современный фотоаппарат с объективом 35мм и качеством 4KHD).



Рисунок 2. Результат

Второй промпт был сфокусирован на создании профессионального и стильного портрета крупным планом (рис.3). В отличие от первого запроса, этот промпт требовал простого черного фона и мягкого, равномерного освещения. Описание включало такие детали, как стильный черный блейзер,

уложенные волнами волосы и гламурный макияж. Цель заключалась в получении чистого, студийного бьюти-изображения.

Промпт: Keep the same person away from the uploaded photo, do not change face or identity. Create a professional and stylish close-up portrait. She is standing against a simple black background. She is wearing a stylish black blazer. Her long hair is styled in loose waves and shines slightly. She has glamorous make-up, including long lashes, a bold eyeliner and nude coloured lips. She is wearing gold earrings and two thin necklaces. Looking at camera, do not wear black gloves, this hold in the left hand a white mobile phone tilted right hand. The lighting is soft and even, with a white ring of light visible behind it, creating a circular outline and illuminating it softly. Professional studio lighting, beauty photography, ultra detailed (На загруженной фотографии не должен быть изображен один и тот же человек, не меняйте лицо или личность. Создайте профессиональный и стильный портрет крупным планом. Она стоит на простом черном фоне. На ней стильный черный блейзер. Ее длинные волосы уложены свободными волнами и слегка блестят. У нее гламурный макияж, в том числе длинные ресницы, яркая подводка для глаз и губы телесного цвета. На ней золотые серьги и два тонких ожерелья. Глядя в камеру, не надевайте черные перчатки, для этого держите в левой руке белый мобильный телефон, наклонив правую руку. Освещение мягкое и равномерное, за ним видно белое кольцо света, создающее круговой контур и мягко освещающее его. Профессиональное студийное освещение, бьюти-фотосъемка с высокой детализацией).

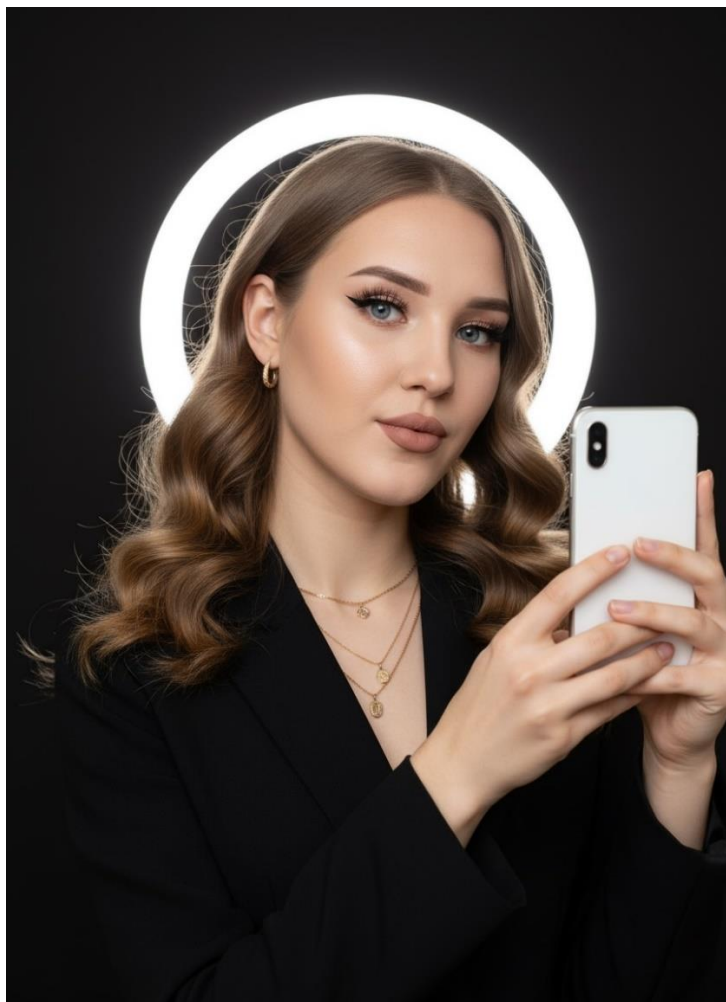


Рисунок 3. Результат

Третий промпт предлагал более модный и интригующий образ (рис.4). Ключевым элементом здесь стали черные кожаные перчатки, хаотично расположенные на лице модели. Запрос требовал высокой детализации текстуры кожи и перчаток, а также использования массивных серебряных украшений для создания ощущения роскоши. Мягкое, но драматичное освещение должно было подчеркнуть контраст между кожей и глянцевой кожей на темном фоне.

Промпт: A fashionable portrait of a woman, close-up, with black leather gloves randomly placed on her face, partially obscuring it. The gloves create an intriguing composition, with one hand covering part of the face and the other resting on the opposite side, subtly framing the model's delicate features. The model's lips are visible, slightly parted, and her eyes are open, with her head slightly tilted forward. Several strands of hair fall freely on her face, adding a sense of movement. The skin texture is highly detailed, revealing the pores of her face. The gloves are shiny and tight fitting, with intricate details. Silver massive rings are visible on the fingers, enhancing the sense of luxury, and silver bracelets wrap around the wrists. The soft yet dramatic lighting highlights the texture of the gloves and leather, creating a striking contrast against the dark, gloomy background. The overall atmosphere is intimate, elegant, and sophisticated (Модный портрет женщины,

крупным планом, черные кожаные перчатки хаотично расположены на ее лице, частично заслоняя его. Перчатки создают интригующую композицию: одна рука закрывает часть лица, а другая лежит на противоположной стороне, тонко обрамляя нежные черты модели. Губы модели видны, слегка приоткрыты, глаза открыты, голова слегка наклонена вперед. Несколько прядей волос свободно падают на ее лицо, добавляя ощущение движения. Текстура кожи очень детализирована, демонстрируя поры ее лица. Перчатки блестящие и плотно прилегающие, с тонкими, замысловатыми деталями. Серебряные массивные кольца видны на пальцах, усиливая ощущение роскоши, а серебряные браслеты обхватывают запястья. Мягкое, но драматичное освещение подчеркивает текстуру перчаток и кожи, создавая поразительный контраст на темном, мрачном фоне. Общая атмосфера интимная, элегантная и изысканная).



Рисунок 4. Результат

Четвертый промпт был нацелен на создание вертикального портрета с кинематографическим освещением и интенсивным контрастом. Съемка под низким углом была призвана подчеркнуть линию подбородка и шеи, создавая

ощущение спокойного доминирования. Яркий малиново-красный фон контрастировал со светлой кожей модели и темной одеждой (черный костюм и рубашка), а направленное освещение создавало резкие тени, подчеркивая структуру лица.

Промпт: Create a Vertical portrait shot in 1080x1920 format, characterized by stark cinematic lighting and intense contrast. Captured with a slightly low, upward-facing angle that dramatizes the subject's jawline and neck, the composition evokes quiet dominance and sculptural elegance. The background is a deep, saturated crimson red, creating a bold visual clash with the model's luminous skin and dark wardrobe. Lighting is tightly directional, casting warm golden highlights on one side of the face while plunging the other into velvety shadow, emphasizing bone structure with almost architectural precision. Wearing smart black suit and black shirt. Shooting half top body (Создайте вертикальный портретный снимок в формате 1080x1920, отличающийся ярким кинематографическим освещением и интенсивным контрастом. При съемке под слегка заниженным углом, обращенным вверх, что подчеркивает линию подбородка и шеи объекта съемки, композиция создает ощущение спокойной доминанты и скульптурной элегантности. Фон выполнен в глубоком, насыщенном малиново-красном цвете, что создает смелый визуальный контраст со светлой кожей модели и темным гардеробом. Освещение направлено в разные стороны, отбрасывая теплые золотистые блики на одну сторону лица, а другую погружая в бархатистую тень, подчеркивая структуру костей с почти архитектурной точностью. одет в элегантный черный костюм и черную рубашку. съемка верхней части тела).

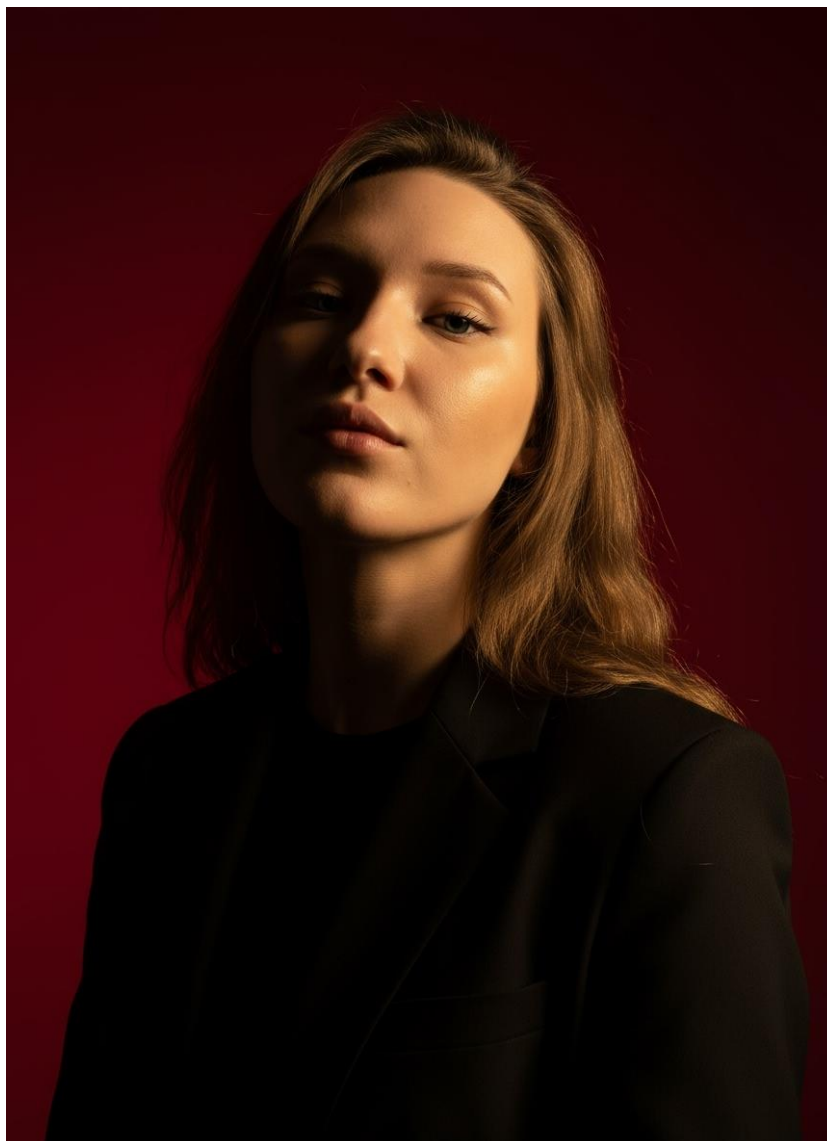


Рисунок 5. Результат

Пятый промпт описывал создание модного портрета с элементами царственности. Модель должна была держать над головой замысловатую черную корону, тень от которой падала на лицо, создавая загадочный эффект. Освещение, направленное сверху, акцентировало корону. Образ дополняло нежное черное кружевное платье с высокой шеей на минималистичном темном фоне.

Промпт: Create a fashionable portrait of a woman standing directly in front of the camera. She holds an intricately detailed black crown above her head. The lighting falls from above, highlighting the crown and creating a dramatic high-contrast effect. The dark shadow of the crown on her face adds to the sense of mystery and elegance. Her dress is a delicate black lace with a high neck, and the frame ends at her shoulders. The background remains dark and minimalist, enhancing the moody and regal atmosphere of the scene (Создай модный портрет женщины, стоящей прямо перед камерой. Она держит замысловато детализированную черную корону над головой. Освещение падает сверху, подчеркивая корону и создавая драматический эффект высокой

контрастности. Темная тень от короны на ее лице усиливает ощущение таинственности и элегантности. Ее платье представляет собой нежное черное кружево с высокой шеей, а рама обрывается на ее плечах. Фон остается темным и минималистским, усиливая угрюмую и царственную атмосферу сцены).

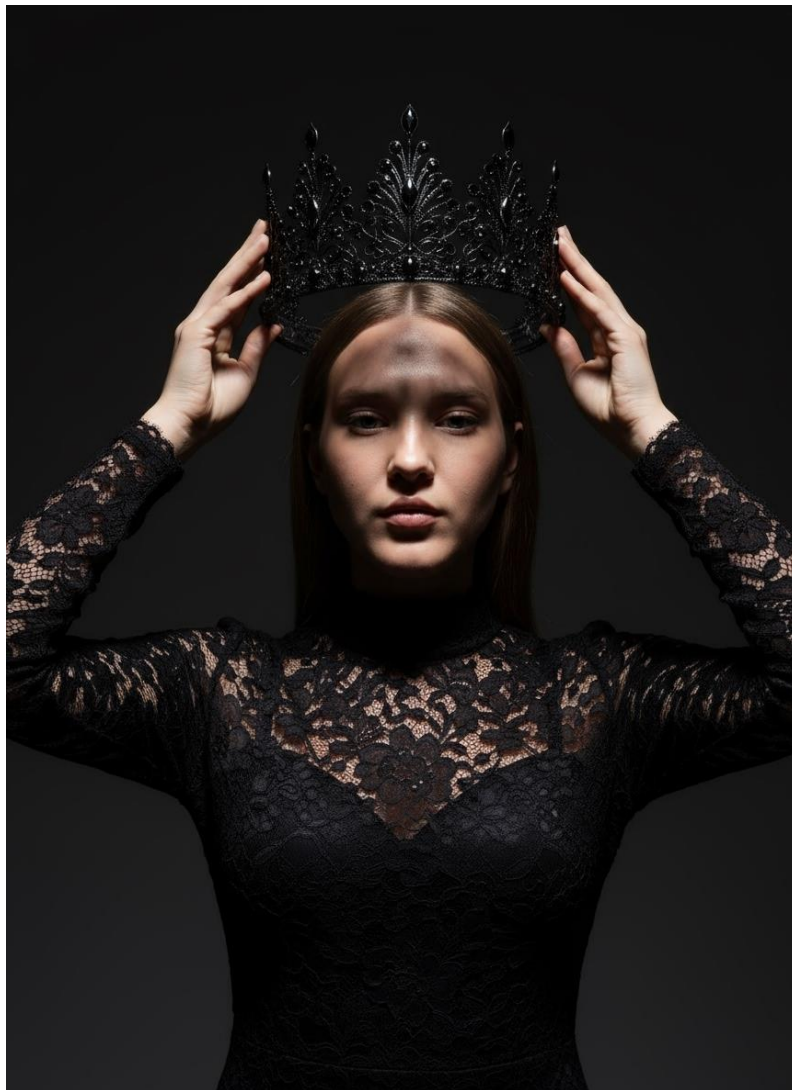


Рисунок 6. Результат

4 Выводы

Проведенное исследование демонстрирует, что текстовые промты являются эффективным инструментом для целенаправленного управления генерацией изображений мультимодальной нейросетью Perplexity. Анализ результатов показывает, что нейросеть способна точно интерпретировать не только конкретные описательные детали (одежда, фон, ракурс), но и абстрактные стилистические концепции, такие как «неон-нуар», «кинематографичность» и «элегантность».

Библиографический список

1. Ильинская Е.В., Гольшева Е.Н., Медведев А.А., Масалитин Н.С. Применение генеративно-состязательных 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 // Научные 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-состязательных 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.