

Генеративный синтез фотореалистичного контента на базе мультимодальной платформы Perplexity: методика и результаты

Пайдершев Алексей Константинович

Приамурский государственный университет имени Шолом-Алейхема

Студент

Аннотация

Цель настоящего исследования заключается в демонстрации функциональных возможностей мультимодальной нейросетевой платформы Perplexity, применяемой для поэтапной генерации фотореалистичных изображений на основе совмещения визуальных референсов и детализированных текстовых промтов. Данный подход обеспечивает целостность композиционных, стилистических и эмоционально-смысловых характеристик визуальных решений в рамках единой нарративной структуры. Выбор платформы Perplexity обусловлен её высокой эффективностью в синтезе фотореалистичных изображений путём интеграции текстовых и визуальных данных. Реализация гибридного метода позволяет, с одной стороны, достоверно воспроизводить индивидуальные особенности визуализируемых объектов, а с другой — осуществлять управляемую модификацию художественных параметров композиции, освещения и стилистики в соответствии с заданными исследовательскими и творческими целями.

Ключевые слова: изображение, промт, нейросеть, Perplexity.

Generative synthesis of photorealistic content based on the Perplexity multimodal platform: methodology and results

Paysershev Alexey Konstantinovich

Sholom-Aleichem Priamurskiy State University

Student

Abstract

The purpose of this study is to demonstrate the functionality of the Perplexity multimodal neural network platform, which is used to generate photorealistic images and video content in a step-by-step manner by combining visual references and detailed text prompts. This approach ensures the integrity of the compositional, stylistic, and emotional-semantic characteristics of the visual solutions within a single narrative structure. The choice of the Perplexity platform is based on its high efficiency in synthesizing photorealistic images by integrating text and visual data. The implementation of the hybrid method allows, on the one hand, to faithfully reproduce the individual features of the visualized objects, and on the other hand, to control the modification of the artistic parameters of the composition, lighting, and style in accordance with the specified research and creative goals.

Keywords: image, prompt, neural network, Perplexity.

1 Введение

1.1 Актуальность

Актуальность данного исследования определяется динамичным развитием мультимодальных нейросетевых технологий, радикально изменяющих подходы к созданию визуального контента в таких областях, как реклама, PR и образование. В условиях интенсификации цифровых коммуникаций наблюдается устойчивый рост потребности в фотореалистичных изображениях, адаптированных под специфику бренда и обладающих выраженным эмоциональным воздействием на аудиторию. Данный спрос существенно превосходит возможности традиционных методов фотосъёмки, характеризующихся значительными временными и ресурсными затратами.

1.2 Обзор исследований

Е.В. Ильинская, Е.Н. Голышева, А.А. Медведев, Н.С. Масалитин описали применение генеративно-согласованных 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-согласованные 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 заключается в демонстрации функциональных возможностей мультимодальной нейросетевой платформы Perplexity, применяемой для поэтапной генерации фотореалистичных изображений на основе совмещения визуальных референсов и детализированных текстовых промтов.

2 Материалы и методы

В качестве основного инструментария исследования использовалась мультимодальная нейросетевая платформа Perplexity, реализующая технологию генеративного синтеза изображений посредством интеграции визуальных референсов (исходных фотографий) и текстовых промтов. Такой подход обеспечивает высокий уровень точности в управлении параметрами композиции, стилистики и визуальных характеристик итогового изображения.

3 Результаты

В процессе первого этапа генерации изображений на платформе Perplexity в систему было внесено референсное изображение (рис. 1) —

молодого человека, которое использовалось для формирования стилистических и композиционных особенностей финальных изображений.



Рисунок 1. Референс

С помощью детально проработанного промта был сгенерирован фотореалистичный портрет в формате 9:16 и разрешении 8К, на котором изображён тот же молодой человек в чёрном войлочном пальто на бордовом фоне с ярко выраженными чертами лица и шеей. В кадре создаётся напряжение благодаря чёткому кадрированию и близкому расположению камеры. Изображение имеет чёткое кинематографическое освещение и сильный контраст. (рис. 2)

Промт: Use this photo and transform it into a vertical portrait in a 9:16 format, characterized by sharp cinematic lighting and strong contrast. The shot is taken from a slightly low, upward-facing angle, emphasizing the model's jawline and neck, creating a sense of quiet dominance and sculptural elegance. The background is a rich, dark crimson red, creating a bold visual contrast with the model's luminous skin and dark wardrobe. The lighting is tightly directed, casting warm golden

highlights on one side of the face, while plunging the other side into velvety shadow, highlighting the bone structure with almost architectural precision.

The model wears a heavy wool or felt overcoat, its texture sharply defined against the smooth, dewy glow of the skin. Minimal retouching preserves skin texture and slight imperfections, adding realism. Editorial tension is created through tight cropping, tonal control, and the almost suffocating proximity of the camera. There are no props or accessories; the visual impact is created purely through light, shadow, color saturation, and posture—evoking high fashion, contemporary isolation, and hyper-modern masculinity. (Используйте эту фотографию и превратите ее в вертикальный портрет в формате 9:16, отличающийся четким кинематографическим освещением и сильным контрастом. Снимок сделан под немного низким углом, обращенным вверх, что подчеркивает линию подбородка и шеи модели, создавая ощущение спокойного доминирования и скульптурной элегантности. Фон выполнен в насыщенных темно-малиновых тонах, что создает яркий визуальный контраст со светлой кожей модели и темным гардеробом. Направленный свет отбрасывает теплые золотистые блики на одну сторону лица, в то время как другая сторона погружается в бархатистую тень, подчеркивая строение костей с почти архитектурной точностью.

Модель одета в пальто из плотной шерсти или войлока, текстура которого четко выделяется на фоне гладкой, влажной от росы кожи. Минимальная ретушь сохраняет текстуру кожи и небольшие дефекты, добавляя реалистичности. Напряжение в кадре создается благодаря четкому кадрированию, регулировке тона и почти удушающей близости камеры. Здесь нет реквизита или аксессуаров)

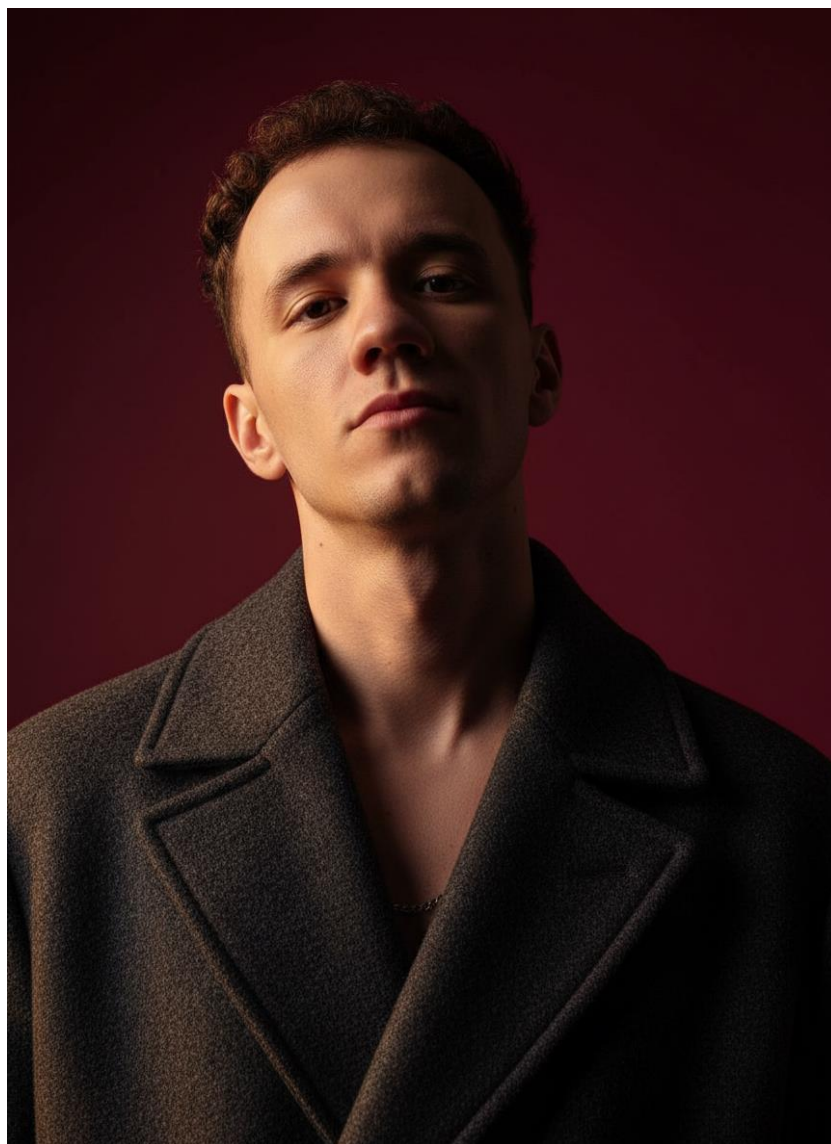


Рисунок 2. Результат

Второй промт позволил создать ещё один фотореалистичный портрет. На нём изображён всё тот же молодой человек в той же позе, но вместо пальто на нём надет чёрный свитшот, а фон не ретуширован и имеет малиновый цвет. Камера располагается ещё ближе, свет изменён на более тёплый, что ещё сильнее выделяет черты лица. Изображение выполнено в разрешении 8K с кинематографичным реализмом. (рис. 3)

Промт: Вертикальный портрет, снятый в масштабе 9:16 с низкого ракурса, позволяет запечатлеть скульптурную четкость скул и ключиц модели. Кинематографическое освещение создает на лице теплые золотистые тона слева, в то время как правое лицо растворяется в глубокой темно-красной тени. Модель одета в черное платье с высоким воротом, выделяющееся на насыщенном, бархатистом малиновом фоне, который подчеркивает сияние ее кожи. (Вертикальный портрет, снятый в масштабе 9:16 с низкого ракурса, позволяет запечатлеть скульптурную четкость скул и ключиц модели. Кинематографическое освещение создает на лице теплые золотистые тона слева, в то время как правое лицо растворяется в глубокой темно-красной тени.

Модель одета в черное платье с высоким воротом, выделяющееся на насыщенном, бархатистом малиновом фоне, который подчеркивает сияние ее кожи.)

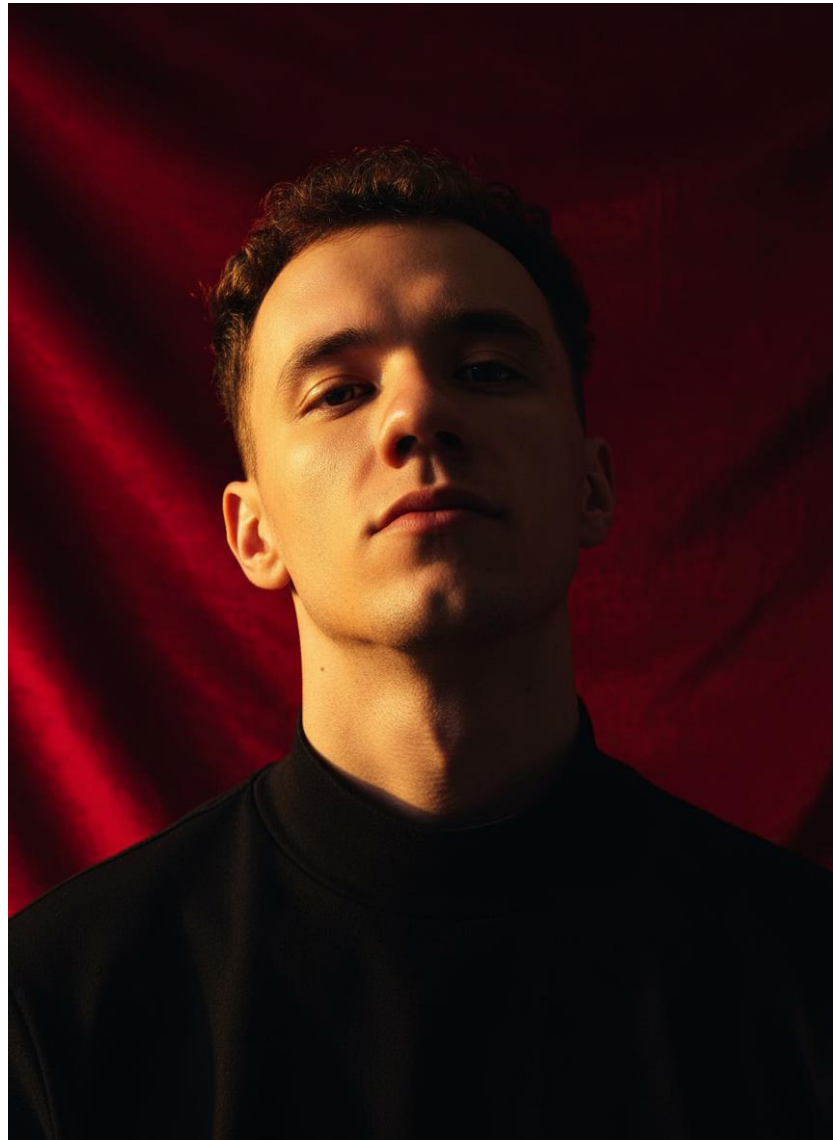


Рисунок 3. Результат

С помощью третьего промта был создан ещё один фотореалистичный вертикальный портрет с разрешением 8K. Модель контрастирует с тёмно-малиновым фоном благодаря чёрной блестящей рубашке и правильно подобранного света. Одна сторона лица сияет янтарными бликами, другая скрыта тенью. (рис. 4)

Промт: A dramatic vertical portrait in 9:16 ratio, photographed from below to exaggerate the model's neck and jawline. Sharp chiaroscuro lighting carves the face—one side glowing with amber highlights, the other swallowed by shadow. The deep crimson backdrop acts as a visual anchor, contrasting the cool elegance of the model's dark attire and radiant complexion. (Эффектный вертикальный портрет в соотношении 9:16, снятый снизу, чтобы подчеркнуть шею и линию подбородка модели. Резкая светотень подчеркивает лицо — одна сторона

сияет янтарными бликами, другая скрыта тенью. Темно-малиновый фон служит визуальным акцентом, контрастирующим с холодной элегантностью темного наряда модели и сияющим цветом ее лица.)



Рисунок 4. Результат

С помощью четвертого промта был создан кинематографический вертикальный кадр, который освещен контрастным светом: одна сторона лица залита теплым золотом, другая скрыта в плотной синевато-черной тени. Небольшой угол съемки подчеркивает объем и присутствие. На фоне глубокого малинового цвета кожа модели выглядит светящейся, а черные наряды сливаются с темнотой, оставляя в фокусе только черты лица. (рис. 5)

Промт: Cinematic vertical portrait (9:16) with bold directional lighting casting half the face in golden warmth, the other in inky shadow. Shot from a subtle low angle to emphasize structure and presence. The model's skin appears luminous against a saturated crimson background, their black outfit dissolving into the darkness to focus solely on facial architecture. (Кинематографический вертикальный портрет (9:16) с ярким направленным освещением,

отбрасывающим на половину лица золотистое тепло, а на другую - чернильную тень. Снимок сделан под небольшим углом, чтобы подчеркнуть структуру и присутствие. Кожа моделей кажется сияющей на насыщенном малиновом фоне, а их черный наряд растворяется в темноте, подчеркивая исключительно архитектуру лица.)



Рисунок 5. Результат

Следующий промт дал нам тёмный вертикальный портрет, который раскрывает фигуру одной модели в густом, объемном свете. Малиновый фон разлит, как расплавленный бархат, а одиночный теплый поток света обнимает левую половину лица, подчеркивая четкие линии и благородство формы, увиденные снизу вверх. Лаконичная экспрессия и строгий черный наряд усиливают ощущение тихого напряжения и непоказной силы. (рис. 6)

Промт: Moody 9:16 vertical composition featuring a lone model under tight, sculptural lighting. The crimson background envelops the frame like liquid velvet, while a single warm beam caresses the left side of the face, accentuating bone structure from a low, reverent angle. Quiet intensity radiates through restrained

expression and monochromatic wardrobe. (Мрачная вертикальная композиция в 9:16, изображающая одинокую модель при плотном, скульптурном освещении. Малиновый фон обволакивает кадр, словно жидкий бархат, в то время как одинокий теплый луч ласкает левую половину лица, подчеркивая структуру костей под низким, благоговейным углом. Сдержанная экспрессия и монохромный гардероб подчеркивают спокойную напряженность.)



Рисунок 6. Результат

Шестой промт помог создать вертикальный кинематографический портрет, снятый с нижнего ракурса, возвышающий фигуру модели, наделяя её величием и уверенностью. Медовый свет ласкает одну половину лица, в то время как другая растворяется в густой малиновой тени. На фоне глубокого красного сияние кожи и бархатная чернота ткани образуют гармонию вечной элегантности и внутренней силы. (рис.7)

Промт: Vertical cinematic portrait (9:16), captured from below to amplify the model's commanding presence. One side of the face bathed in honeyed light, the other lost in deep crimson-infused shadow. The interplay of luminous skin and matte

black fabric against the rich red backdrop evokes timeless elegance and restrained power. (Вертикальный кинематографический портрет (9:16), снятый снизу, чтобы подчеркнуть внушительный вид модели. Одна сторона лица залита медовым светом, другая скрыта в глубокой малиновой тени. Сочетание сияющей кожи и матовой черной ткани на насыщенном красном фоне создает ощущение неподвластной времени элегантности и сдержанной мощи.)



Рисунок 7. Результат

С помощью седьмого промта получилось создать вертикальный кадр, выдержанный в лаконичном серийном ключе, пронизанный резким, контрастным освещением: диагональный луч делит лицо на стороны — золотистое свечение и холодную тень. Низкий угол взгляда подчеркивает упругое напряжение шеи и четкость очертаний. На фоне плотного малинового пространства модель кажется воплощением молчаливой силы. (рис.8)

Промт: A series-style vertical portrait in 9:16, lit with high-contrast cinematic lighting that slices the face diagonally—golden warmth on the lit plane, cool depth in the shadows. The low angle highlights the neck's tension and jawline's sharpness.

Set against a saturated crimson void, the model exudes silent authority. (Вертикальный портрет в формате 9:16, выполненный в серийном стиле, с высококонтрастным кинематографическим освещением, которое разрезает лицо по диагонали - золотистое тепло на освещенной плоскости, прохладная глубина в тени. Низкий угол обзора подчеркивает напряжение шеи и резкость линии подбородка. На фоне насыщенной малиновой пустоты модель излучает молчаливую властность.)

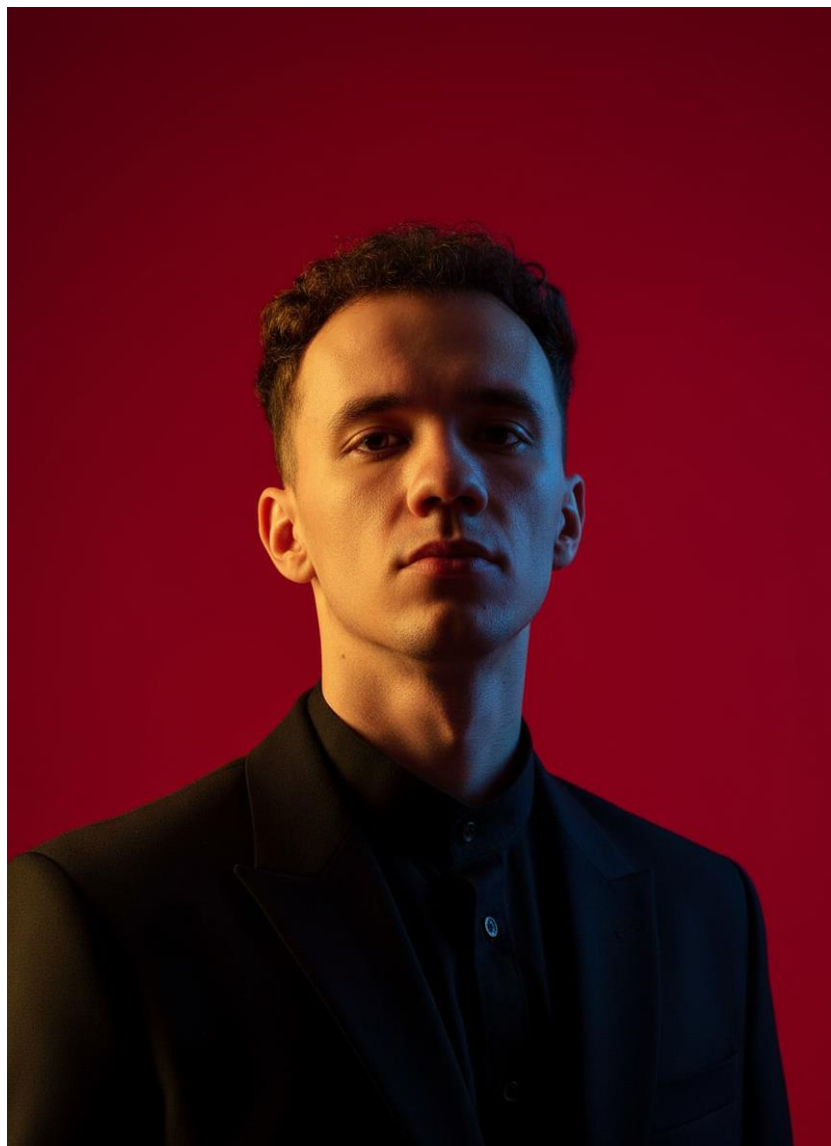


Рисунок 8. Результат

Восьмой промт создаёт вертикальный снимок в пропорции 9:16, сделанный с низкой позиции, который охватывает лишь лицо и плечи. Направленный теплый свет скользит вдоль висков и подбородка, подчеркивая рельеф, а густой малиновый фон тает в черноте ткани, оставляя в центре внимания лицо — вырезанное контрастом, симметрией и тихим напряжением. (рис.9)

Промт: Tightly framed 9:16 vertical shot, emphasizing only face and upper torso from a low vantage point. Warm, directional light grazes the model's temple

and jaw, while the crimson background merges with the dark fabric of their outfit, isolating the face as a study in contrast, geometry, and controlled emotion. (Четко очерченный вертикальный снимок в масштабе 9:16, на котором с низкой точки обзора видны только лицо и верхняя часть туловища. Теплый направленный свет скользит по вискам и подбородку модели, в то время как малиновый фон сливается с темной тканью одежды, выделяя лицо с помощью контрастов, геометрии и контролируемых эмоций.)

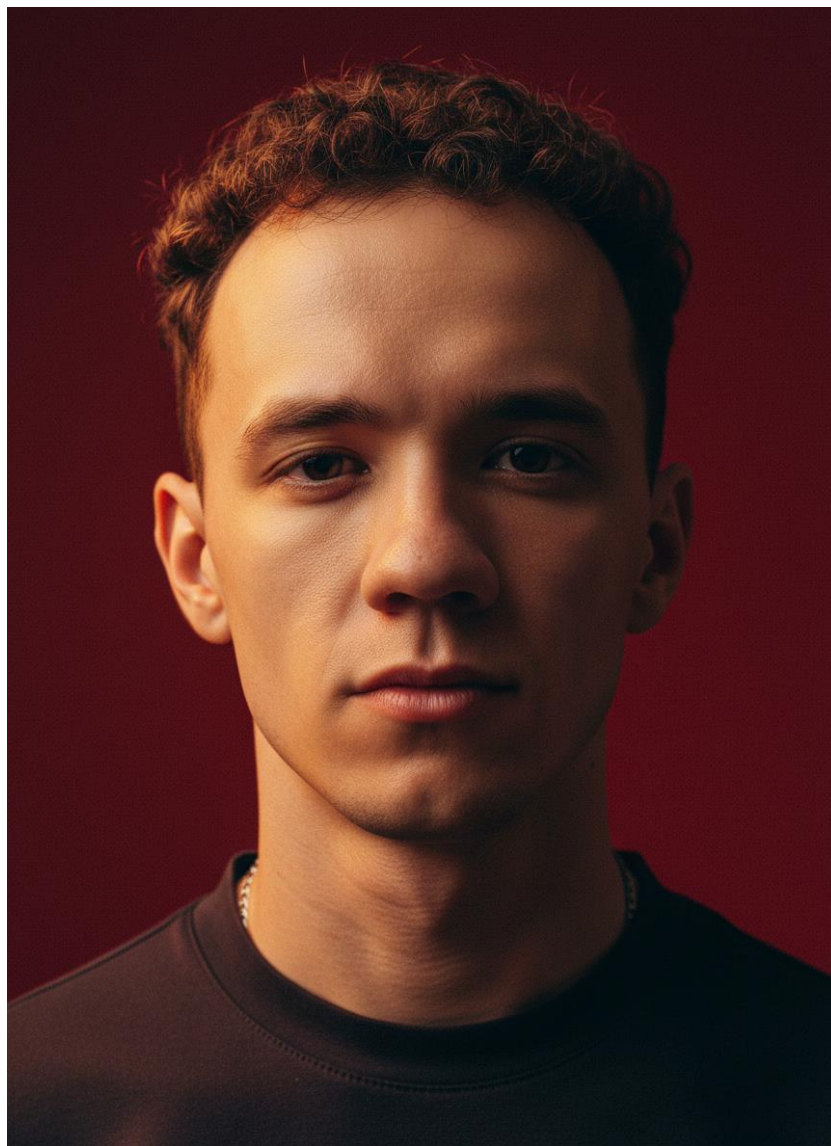


Рисунок 9. Результат

С помощью девятого промта получилось создать вертикальный портрет, где свет, направленный снизу вверх, будто высекает черты модели из камня. Глубокий темно-красный фон обволакивает края кадра, а тонкий золотистый луч мягко очерчивает левый профиль. Черный аскетичный наряд и спокойное, монументальное выражение лица завершают образ, наполненный силой и грацией. (рис.10)

Промт: Architectural portrait in 9:16 ratio—light sculpts the model's features like stone under a low, upward gaze. Crimson darkness wraps the edges of the frame,

while a narrow beam of golden light defines the left profile. The model's minimal black attire and serene expression complete this study in dominance and grace. (Архитектурный портрет в соотношении 9:16 — свет, направленный снизу вверх, словно вылепливает черты модели из камня. Темно-красный цвет окутывает края кадра, в то время как узкий луч золотистого света очерчивает левый профиль. Строгий черный наряд модели и безмятежное выражение лица дополняют это исследование своим доминированием и грацией.)

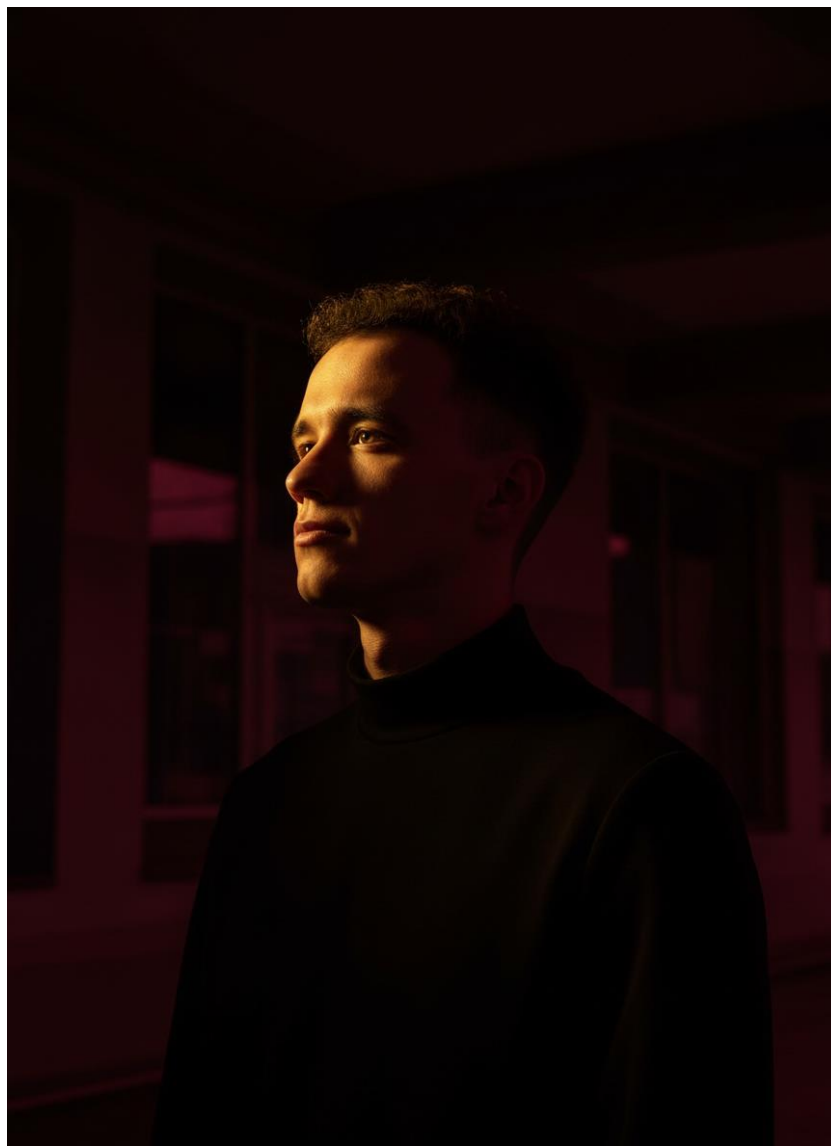


Рисунок 10. Результат

Десятый промт помогает создать вертикальный портрет в формате 9:16, наполненный мягкой игрой света и малиновых полутонов. Позади модели тени перетекают в насыщенный фон, а свет вычерчивает половину лица золотистым теплом, оставляя вторую в полумраке тайны. Низкий вырез словно продолжает линию шеи, создавая статную утонченность, подчеркнутую черным матом ткани и безупречной кожей. (рис. 11)

Промт: Vertical cinematic frame (9:16), where shadow and crimson blend seamlessly behind the model. Lighting is precise and dramatic—half the face glows

with warmth, the other recedes into mystery. The low angle elongates the neck, lending a statuesque quality to the composition, underscored by matte black clothing and flawless skin. (Вертикальный кинематографический кадр (9:16), на котором тени и малиновый цвет плавно сливаются за спиной модели. Освещение точное и эффектное — половина лица излучает тепло, другая часть окутана тайной. Низкий вырез удлиняет шею, придавая композиции статный вид, подчеркнутый матово-черной одеждой и безупречной кожей.)



Рисунок 11. Результат

При помощи последнего промта получается создать высококонтрастный вертикальный портрет, сотканный из темных тонов и малиновых отражений. Глубокий фон будто светится изнутри, а теплый луч, идущий снизу, вылепливает форму модели. Продуманность света и совершенство теней создают гармонию, в которой звучат спокойная сила и утончённая, вечная красота. (рис. 12)

Промт: Moody, high-contrast vertical portrait in 9:16. The crimson background pulses with depth, while a single source of warm light traces the model's

bone structure from below. Every shadow is deliberate, every highlight refined—creating a unified series piece radiating quiet strength and timeless sophistication. (Мрачный, высококонтрастный вертикальный портрет в 9:16. Малиновый фон излучает глубину, а единственный источник теплого света обрисовывает фигуру модели снизу. Каждая тень продумана, каждый блик изыскан — все это создает единое изделие серии, излучающее спокойную силу и неподвластную времени изысканность.)

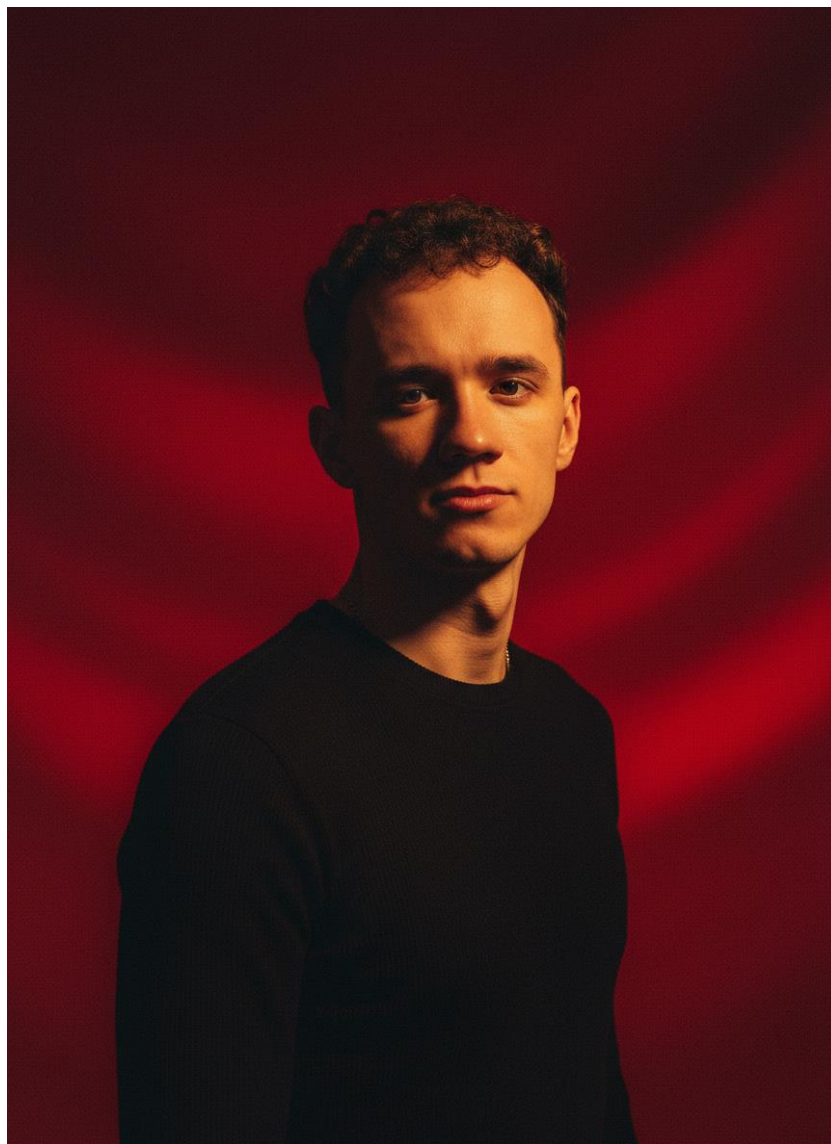


Рисунок 12. Результат

4 Вывод

В результате поэтапного процесса генерации визуального контента на платформе Perplexity, основанного на интеграции визуальных референсов и детализированных текстовых промтов, были получены стилистически согласованные фотореалистичные портреты. Созданные изображения, выполненные в разрешении 8K и обладающие кинематографическим качеством визуализации, свидетельствуют о высокой эффективности предложенной методики.

Проведённый эксперимент подтвердил возможность точного управления композиционными, световыми и пластическими параметрами посредством структурированных текстовых инструкций, что открывает перспективы практического применения технологии в области рекламных и коммуникационных стратегий бренда.

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 // Научное обозрение – технические науки. 2023. №2. С. 10.
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.