

АВТОМАТИЗАЦИЯ РАЗРАБОТКИ ИЗОБРАЖЕНИЙ МНЕМОСХЕМ В ПРОЕКТНОЙ ДОКУМЕНТАЦИИ АСУ ТП ДЛЯ ОБЪЕКТОВ ДОБЫЧИ И ПОДГОТОВКИ НЕФТИ И ГАЗА

Голецихин Д.А.¹, Зебзеев А.Г.²

¹НИ ТПУ, ИШИТР, гр. 8Т11, студент, dag36@tpu.ru

²НИ ТПУ, ИШИТР, к.т.н., доцент ОАР, agz@tpu.ru

АО «ТомскНИПИнефть», Экспертная группа, Главный эксперт по цифровизации

Аннотация

В работе рассмотрена автоматизация разработки изображений мнемосхем в проектной документации АСУ ТП для объектов добычи и подготовки нефти и газа, в основе которой лежит дообучение генеративной нейронной сети, в целях уменьшения временных затрат на создание документации.

Ключевые слова: мнемосхема, генеративная нейронная сеть, дообучение, генерация изображений.

Введение

В нефтегазовой отрасли, прежде чем обустроить объекты для добычи или подготовки нефтегазовой продукции, проектные институты разрабатывают проектную документацию, в которой описывается, как должны работать те или иные объекты. В частности, разрабатывается документ с изображениями того, как должны выглядеть мнемосхемы на автоматизированном рабочем месте (АРМ) оператора. В настоящее время создание этих изображений происходит вручную с использованием SCADA-систем и занимает от 20 минут до 1 часа. Инструментом для решения данной проблемы могут стать стремительно развивающиеся нейронные сети, которые позволяют генерировать изображения высокого качества. В связи с этим целью данной работы является дообучение модели генеративной нейронной сети, способной генерировать изображения мнемосхем по текстовому описанию.

Основная часть

Так как для проектной документации требуется высокое качество изображений, для дообучения была выбрана диффузионная модель нейронной сети, которая благодаря своей архитектуре позволяет достичь необходимого качества. [1]. В данной модели генерация изображений происходит путем итеративного вычитания шума из случайного шума (см. рисунок 1). Это сделано для того, чтобы получать разные результаты при одном и том же запросе. Ее архитектура представлена на рисунке 2.

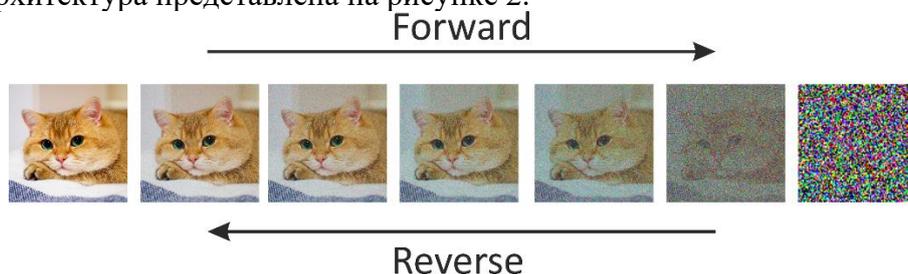


Рис. 1. Прямая и обратная диффузия

Таким образом, чтобы обучить модель, нужно научить ее вычитать шум. Для этого при обучении изображения из датасета, предварительно сжатые в скрытое пространство (latent space) для уменьшения вычислительных потребностей, подвергаются процессу прямой диффузии – итеративному добавлению гауссовского шума к исходному изображению. Далее зашумленное изображение подается в блок Denoising U-net, в котором происходит процесс

обратной диффузии. Этот блок представляет собой энкодер, выделяющий ключевые признаки изображения, сжимая его, и декодер, который восстанавливает изображения до исходных размеров, но с более ярко-выраженными ключевыми признаками. Чтобы избежать потери точности при разжимании изображения, используются связи между слоями энкодера и декодера – skip-connections. Они позволяют декодеру сохранить качество, опираясь на слои при сжатии. Также в блок Denoising U-net подается текстовая подсказка, которая предварительно преобразуется в числовые векторы (conditioning) с помощью текстового энкодера. Эти вектора используются в методе Concat (склеивание числовых векторов с поданными в U-net данными) и слоях перекрестного внимания (Cross-attention), встраиваемые к слоям энкодера и декодера U-net. После этого метод Concat, используемый на начальном этапе вычитания шума, позволяет инициализировать объект, а на следующих шагах включаются слои перекрестного внимания, которые лучше справляются с контролем генерации на основе текста. Используя эти методы, U-net предсказывает шум, который необходимо вычесть для получения «чистого» изображения. Количество итераций в Denoising U-net равно количеству итераций, произведенных при прямой диффузии. После удаления шума, изображение из скрытого пространства разжимается обратно в пиксельное пространство (Pixel Space). Полученное изображение сравнивается с исходным, после чего вычисляется ошибка, на основе которой вычисляются корректирующие значения для весов модели нейронной сети [2, 3].

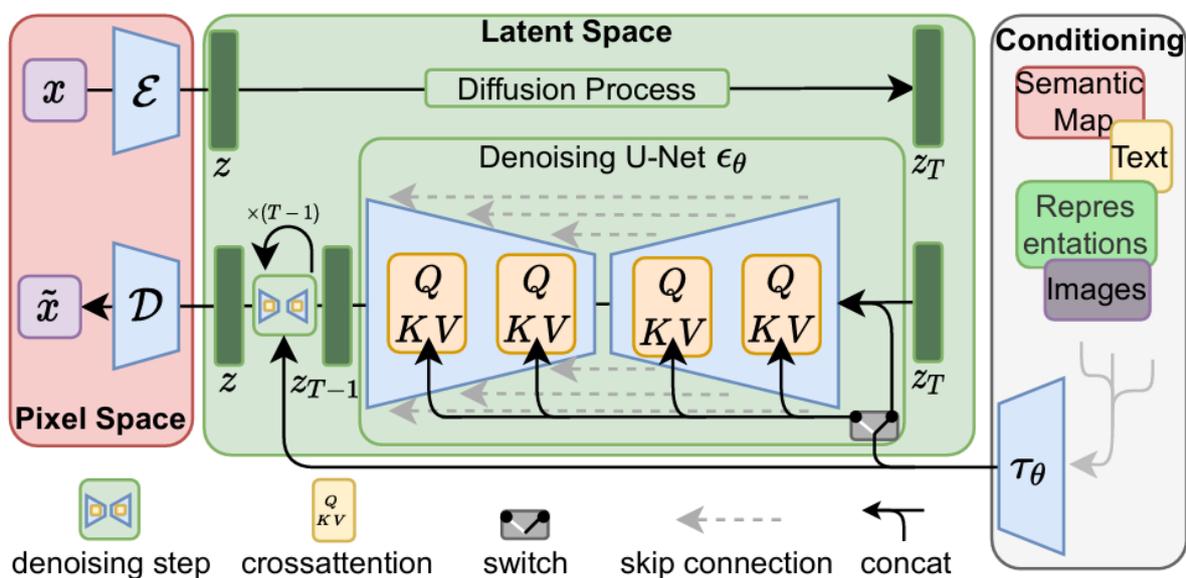


Рис. 2. Архитектура диффузионной модели

Denoising step – этап удаления шума

Cross-attention – слой перекрестного внимания

Switch – переключатель между слоями перекрестного внимания и методом Concat

Skip connection – связь между энкодером и декодером для U-net

Concat – метод используемый для начальной инициализации объекта при расшумлении

Дообучение нейросети – это добавление новых или же изменение основных слоев параметров уже обученной модели нейронной сети, на основе датасета из небольшого количества данных.

Методом дообучения была выбрана LoRA – Low Rank Adaptation (низкоранговая адаптация). Суть данного метода заключается в добавлении в исходную модель слоев с параметрами, обученными на собственном датасете, при этом исходная модель остается неизменной. Преимуществами этого метода дообучения являются небольшой размер за счет хранения параметров в двух матрицах меньшего ранга [4].

Также были рассмотрены два других метода дообучения: DreamBooth и Textual Inversion (текстовая инверсия). Однако они не были выбраны, так как DreamBooth изменяет слои базовой модели, и на выходе дообучения получается модель с большим размером. А текстовая инверсия не подходит, так как ее принцип заключается в дообучении текстового энкодера, который впоследствии будет стараться найти из весов основной модели те, которые схожи с датасетом при дообучении. Но, так как использование изображений мнемосхем при обучении основных моделей маловероятно из-за их узкой направленности и недоступности в общем доступе в большом количестве, то качество получаемых генераций снизится.

Следующим этапом было создание датасета, состоящего из изображений мнемосхем для установок по подготовке нефти и газа: сепараторы, осушители. Датасет из 14 установок был составлен в SCADA-системе, после чего были созданы скриншоты размерностью 512×512, к которым было составлено текстовое описание на английском языке. Пример из датасета представлен ниже.

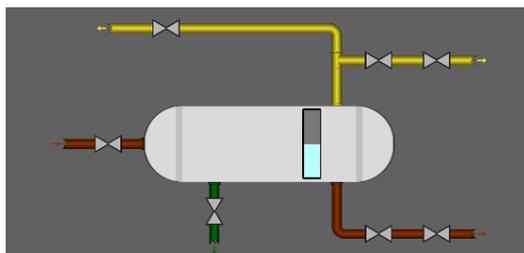


Рис. 3. Пример изображения из датасета

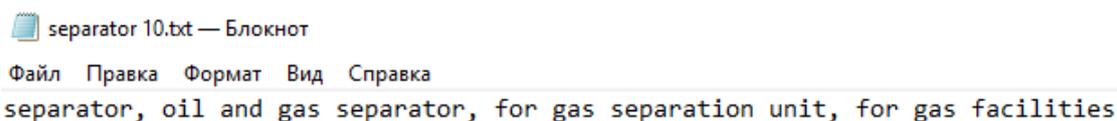


Рис. 4. Пример текстового описания к изображению из датасета

После проведения подготовительных мероприятий в заранее установленном пользовательском интерфейсе, предназначенном для работы с генеративными нейросетями, было проведено дообучение базовой модели Stable Diffusion v1.5. В ходе 4 попыток дообучения (график потерь при последней попытке дообучения приведен на рисунке 5), в процессе которых происходили корректировки обучающих параметров, были получены результаты, сравнимые с реальными изображениями мнемосхем, составленных вручную в SCADA-системе. Результат генерации приведен на рисунке 6. Время генерации занимает около 5 секунд на ПК со следующими характеристиками:

- Процессор Intel Core i9-14900K.
- Материнская плата MSI PRO Z790-A.
- Видеокарта KFA2 GeForce RTX 4070Ti SUPER.
- Оперативная память Kingston Fury Beast Black AMD 64 ГБ × 2 шт.

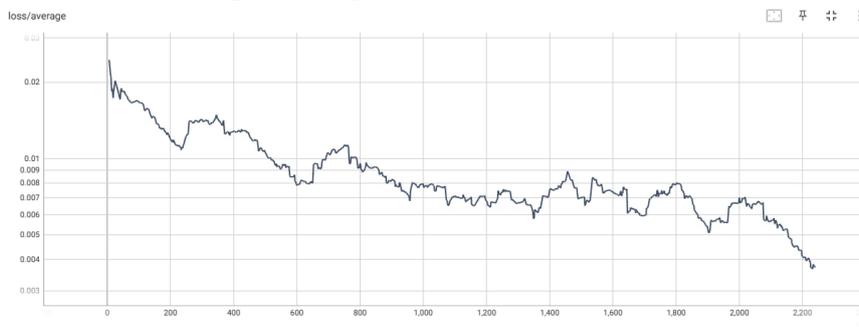


Рис. 5. График потерь при дообучении

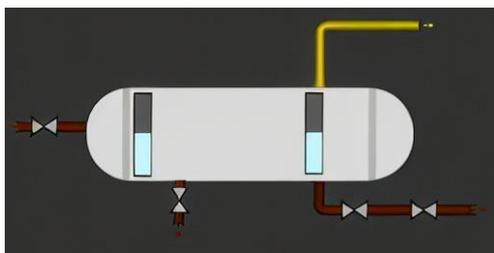


Рис. 6. Результат генерации после дообучения

В ходе реализации данной идеи также было найдено решение для создания более точных генераций – использование ControlNet. Это дополнительная управляющая нейронная сеть, которую можно включить в этап генерации изображений [5]. Принцип ее работы заключается в том, что она позволяет выделить из изображения, поданного на ее вход, признаки, такие как глубина изображения, очертания, поза, положение объекта, а также прямые линии, например, на изображениях интерьера. А далее, выделенный признак может быть использован в генерации изображения [6].

Исходя из возможностей ControlNet и того, что все изображения мнемосхем создаются согласно чертежам технологических схем и функциональных схем автоматизации, можно реализовать генерацию, используя форму, полученную из чертежей с помощью ControlNet. Данный способ должен помочь справиться с проблемой создания более громоздких изображений мнемосхем.

Заключение

В ходе проделанной работы была получена дообученная модель генеративной нейронной сети, позволяющая генерировать изображения мнемосхем объектов подготовки нефти и газа по текстовому описанию на английском языке. Полученные изображения визуально сравнимы с изображениями, созданными вручную. После генерации на изображения остается только расставить поля для измеряемых параметров и подписать для них теги (индивидуальные для каждого проекта).

Так, данная работа может облегчить труд и уменьшить затраты времени для сотрудников отделов проектирования автоматизированных систем.

Планируется следующее развитие работы:

- Улучшение качества модели.
- Применение ControlNet для создания генераций, соответствующих технологическим схемам, на основе которых создается проектная документация.

Список использованных источников

1. Введение в диффузионные модели для генерации изображений [Электронный ресурс]. – URL: habr.com/ru/articles/713076/ (дата обращения: 02.12.2024).
2. Как работают text2image модели? [Электронный ресурс]. – URL: habr.com/ru/articles/711020/ (дата обращения: .07.12.2024).
3. Stable Diffusion: text-to-person [Электронный ресурс]. – URL: habr.com/ru/articles/764700/ (дата обращения: 10.12.2024).
4. Кто же такая это ваша LoRA [Электронный ресурс]. – URL: habr.com/ru/articles/747534/ (дата обращения: 11.12.2024).
5. ControlNet: расширение возможностей нейросетей [Электронный ресурс]. – URL: t-j.ru/controlnet/ (дата обращения: 25.02.2025).
6. Как работает ControlNet. Контролируемая генерация [Электронный ресурс]. – URL: habr.com/ru/companies/ruvds/articles/719348/ (дата обращения: 26.02.2025).