

Stable Diffusion:

Техники и Подходы в генерации

Гайнетдинов А.Ф.

Team Lead, Senior Deep Learning Engineer, VicMan LLC

zirgan.ainur@gmail.com

Применение Генеративных Моделей

Искусство & Дизайн



Создание контента



Основа для обучения
новых моделей



Развлечения



Маршрут доклада

1. Генеративные модели. Диффузия из термодинамики для генерации изображений.
2. Что умеют модели по умолчанию.
3. Обучение моделей новым образам.
4. Расширенные возможности SD

Генеративные Модели

Kandinsky



DALL-E

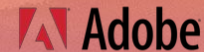


OpenAI

Imagen



Firefly



Stable Diffusion

stability ai

Midjourney

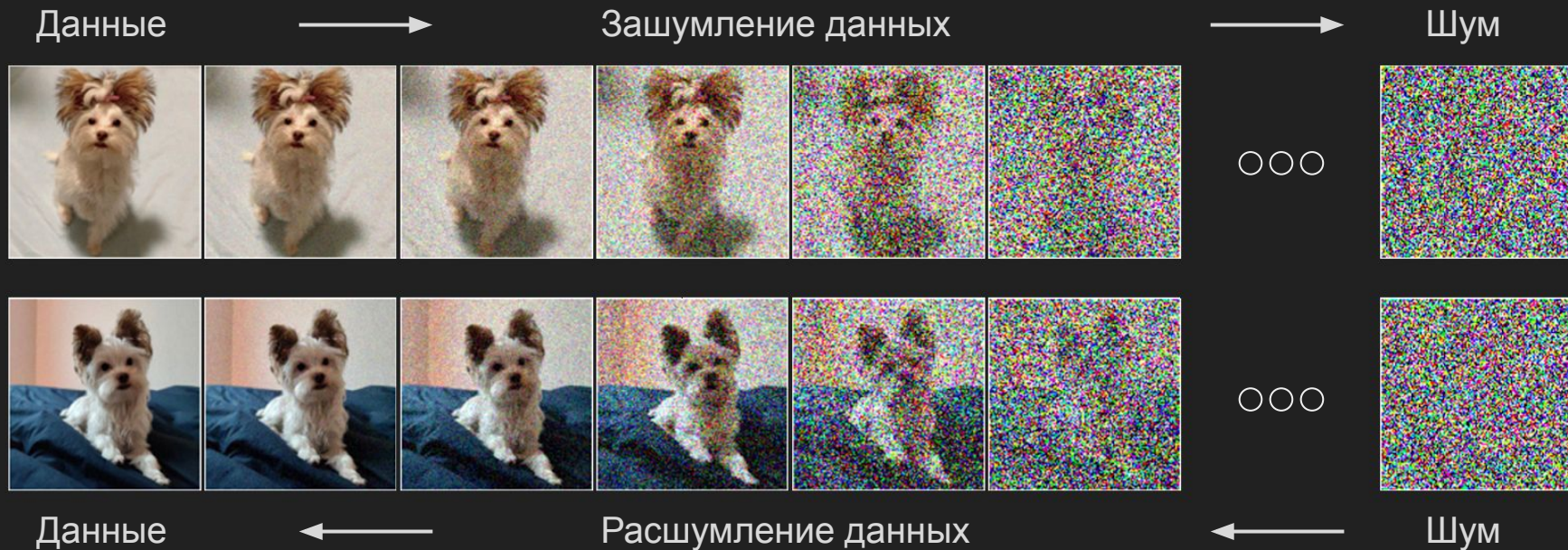


Midjourney

Stable Diffusion versions

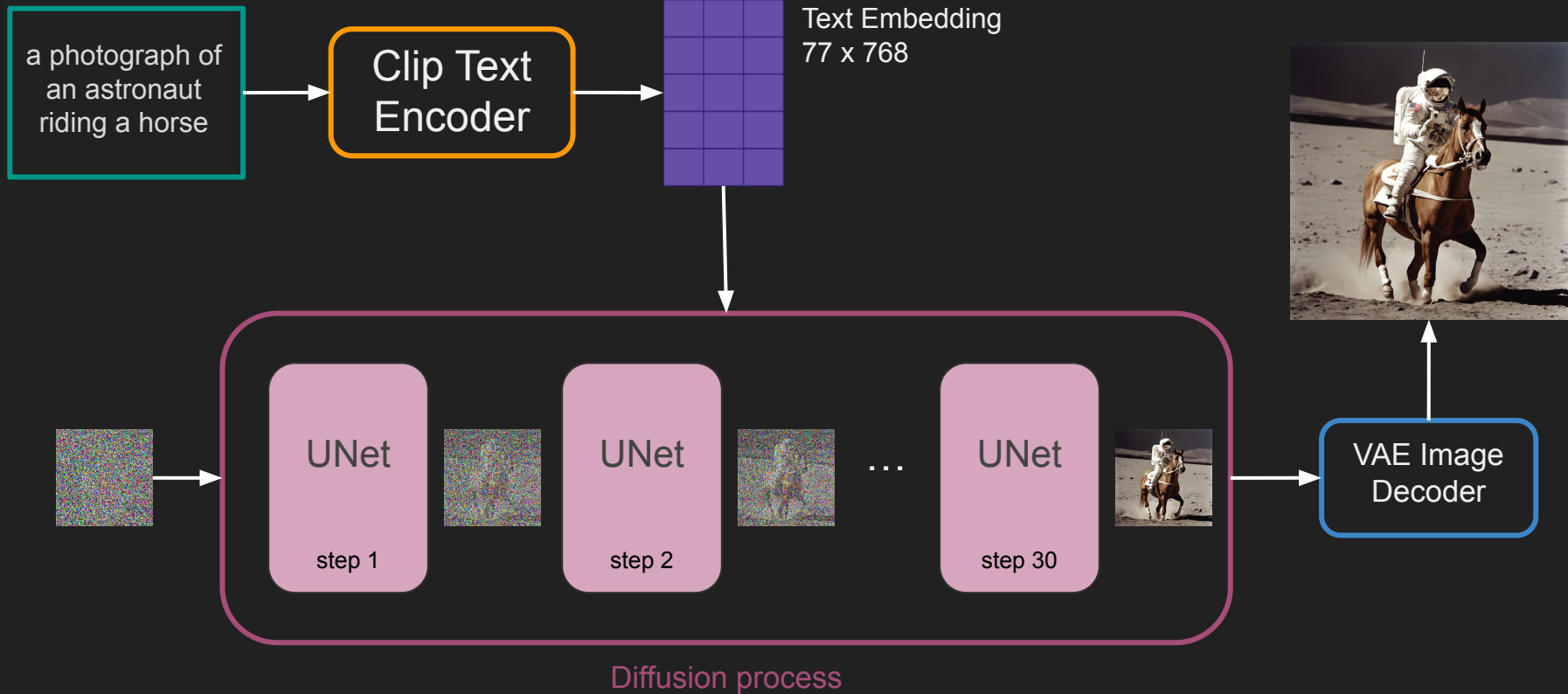
Version number	Release date	Notes
1.1, 1.2, 1.3, 1.4 ^[57]	August 2022	All released by CompVis. There is no "version 1.0". 1.1 gave rise to 1.2, and 1.2 gave rise to both 1.3 and 1.4. ^[58]
1.5 ^[59]	October 2022	Initialized with the weights of 1.2, not 1.4. Released by RunwayML.
2.0 ^[60]	November 2022	Retrained from scratch on a filtered dataset. ^[61]
2.1 ^[62]	December 2022	Initialized with the weights of 2.0.
XL 1.0 ^{[63][21]}	July 2023	The XL 1.0 base model has 3.5 billion parameters, making it around 3.5x larger than previous versions. ^[64]
XL Turbo ^[65]	November 2023	Distilled from XL 1.0 to run in fewer diffusion steps. ^[66]
3.0 ^{[67][22]}	February 2024 (early preview)	A family of models, ranging from 800M to 8B parameters.

Процесс генерации Stable Diffusion



Что умеют модели по умолчанию

Архитектура Stable Diffusion 1.5



Text 2 Image



a cat drinking a pint of beer

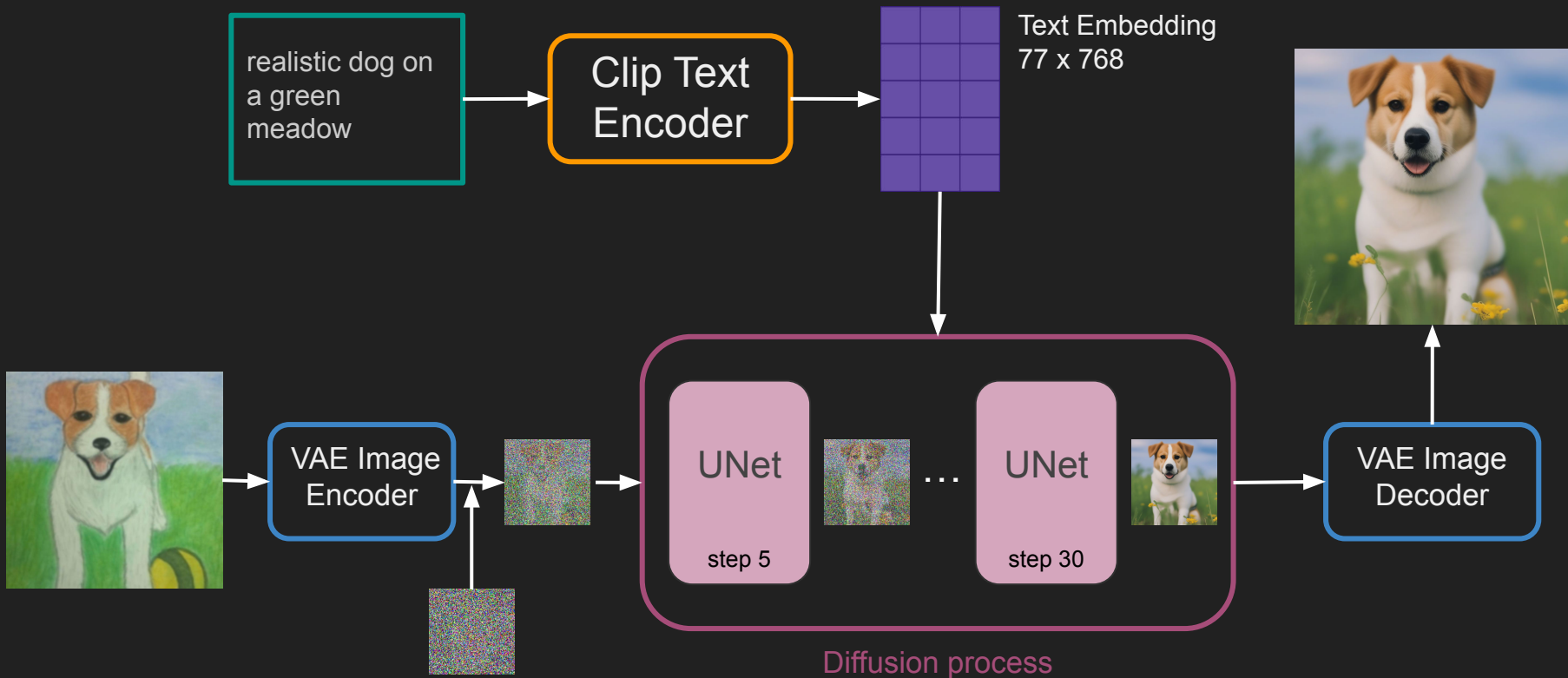


a baby penguin wearing a blue hat, red gloves, green shirt, and yellow pants



a cute robot artist painting on an easel
concept art

Image 2 Image



SDXL Inpainting

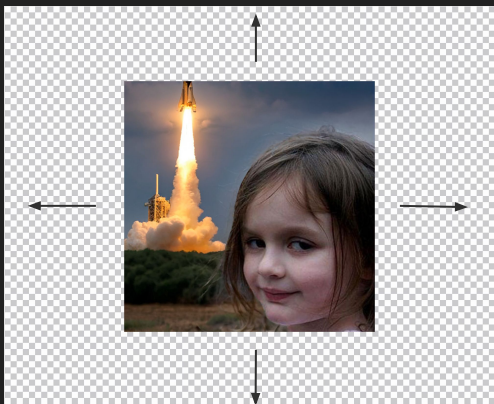


Inpainting mask

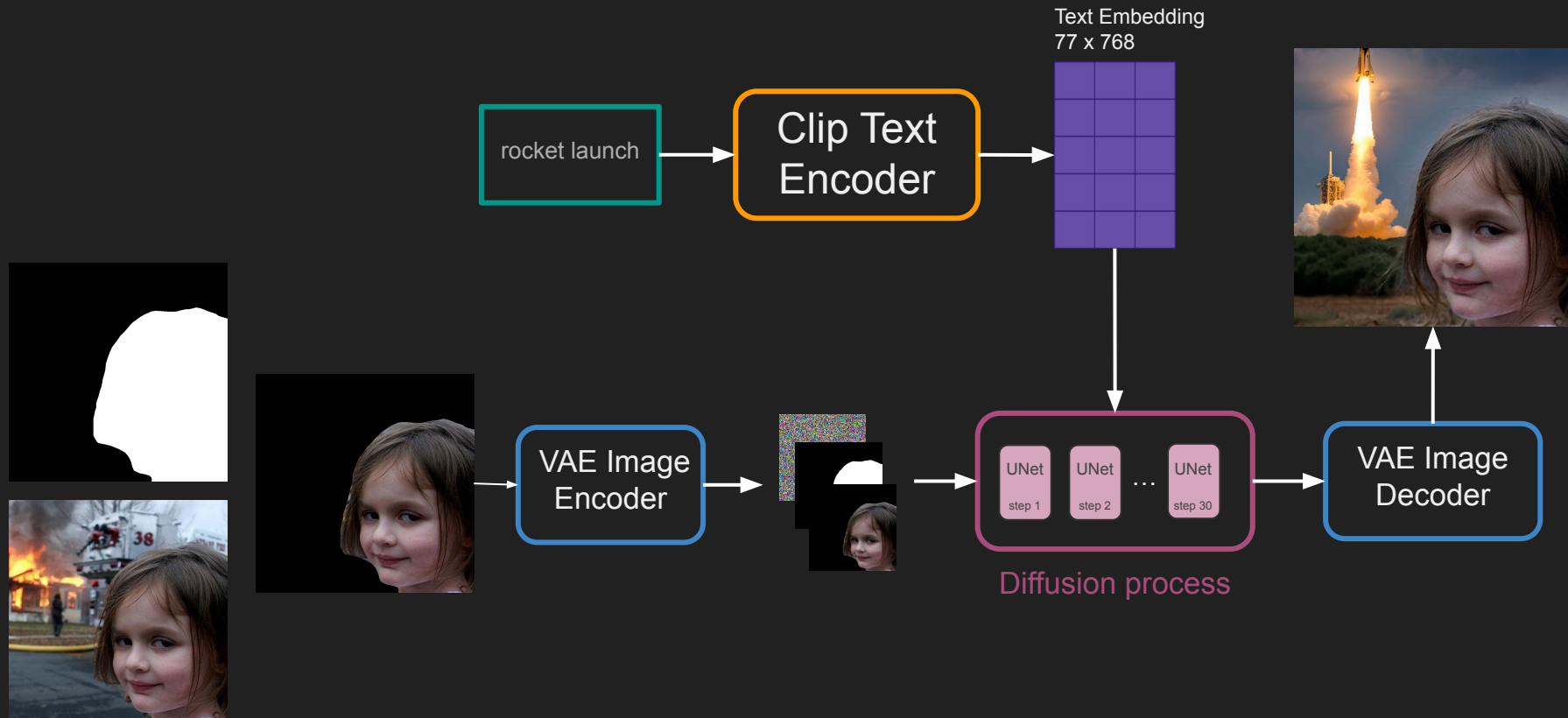


Prompt: rocket launch

SDXL Outpainting



Stable Diffusion Inpainting



Графические интерфейсы

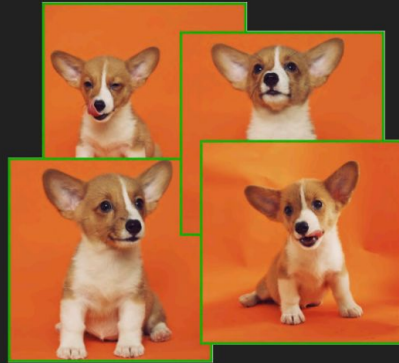
- [Fooocus](#) - простой в использовании, высокое качество, встроенные стили
- [ComfyUI](#) - сложный в использовании, много возможностей, гибкий, понятный код, API
- [AUTOMATIC1111](#) - средний в использовании, много возможностей, API

Обучение моделей новым образам

Fine-tuning

Expression modification (“A [state] [V] dog”)

Input images



Fine-tuning

Input images



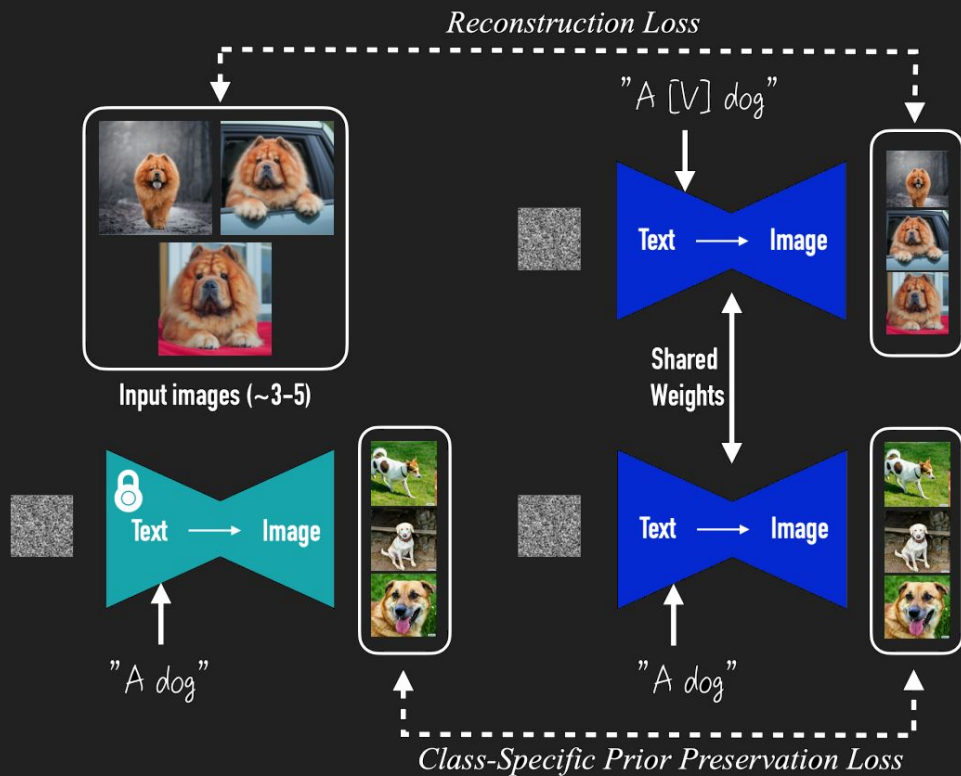
w/o prior-preservation loss



with prior-preservation loss



Fine-tuning (DreamBooth)



Тренируем Lo Stable Diffusion 1.5 вашему лицу

Генерация фото

Фото для обучения



0000.png



0001.png

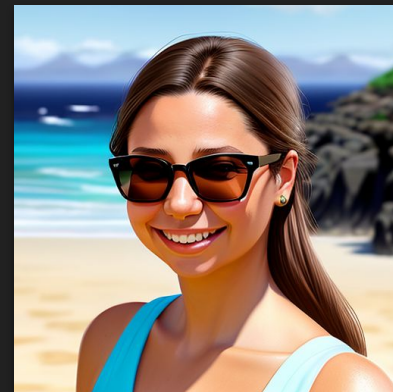


0002.png

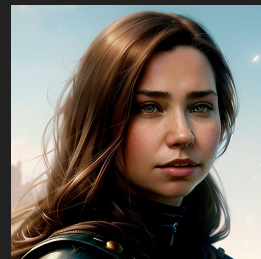
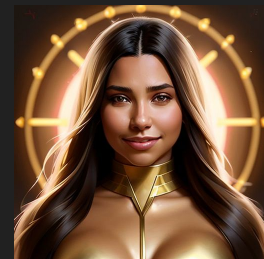
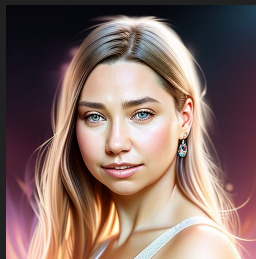
Training prompt: a photo of **zwx** woman
Regularization prompt: a photo of woman



a photo of **zwx** woman, blue skin, yellow eyes, diffuse natural sun lights, Na'vi race Avatar movie, James Cameron, Toruk Makto, forest background

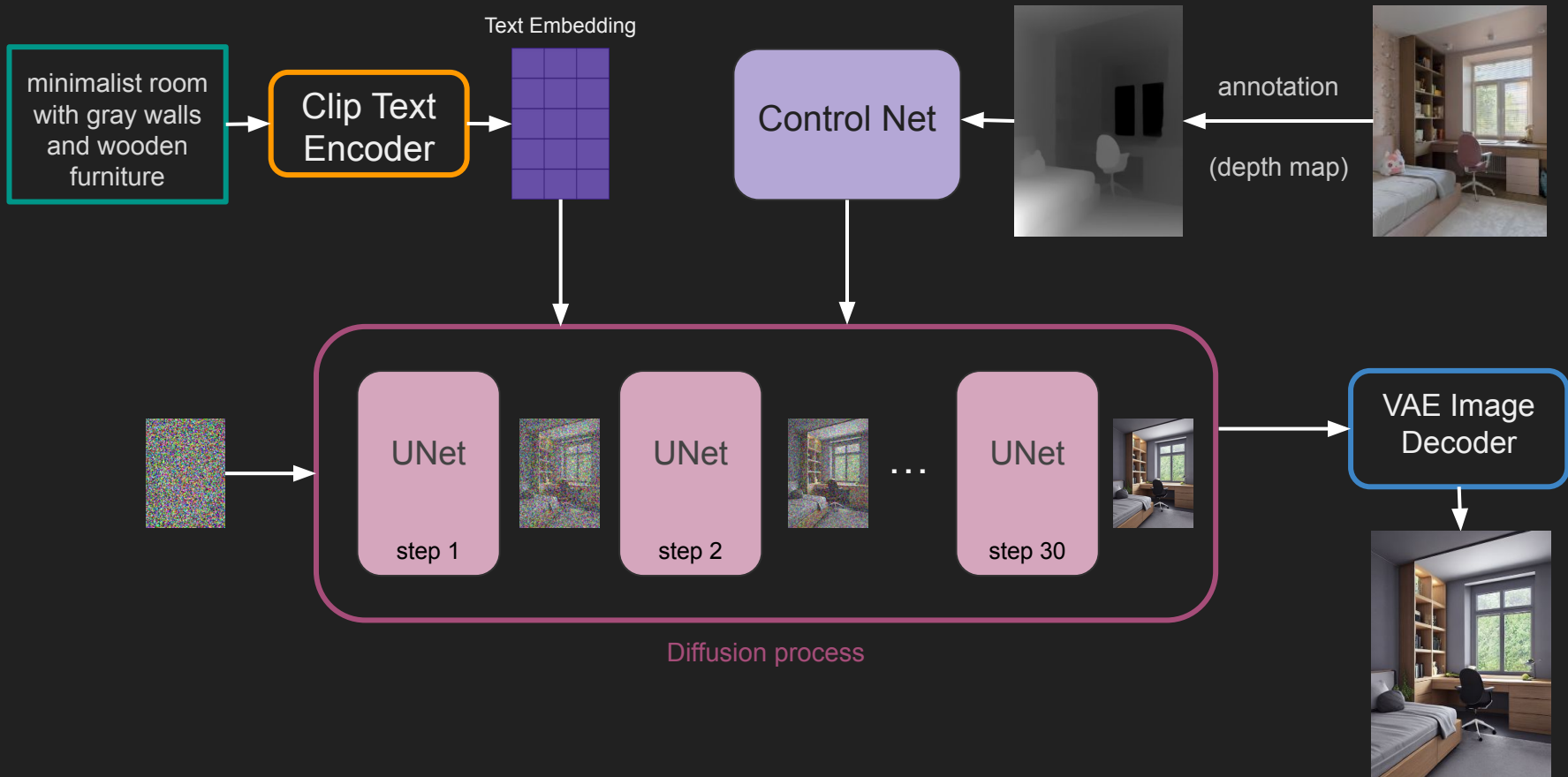


a photo of **zwx** woman, happy joy, slight smile, diffuse natural sun lights, summer mood, sunglasses, sun on face, sea background



Расширенные возможности Stable Diffusion

Control Net



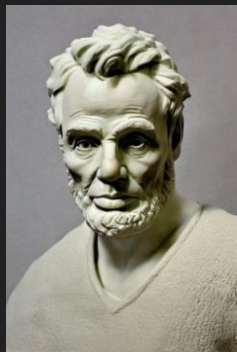
Control Net Annotators



Input



Normal Map



plaster statue of
Abraham Lincoln



Input



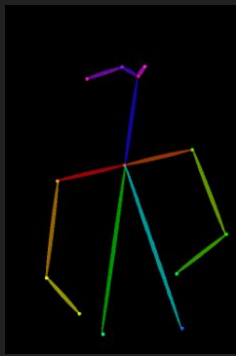
Canny



cute dog



Input



OpenPose



chef in kitchen



Input



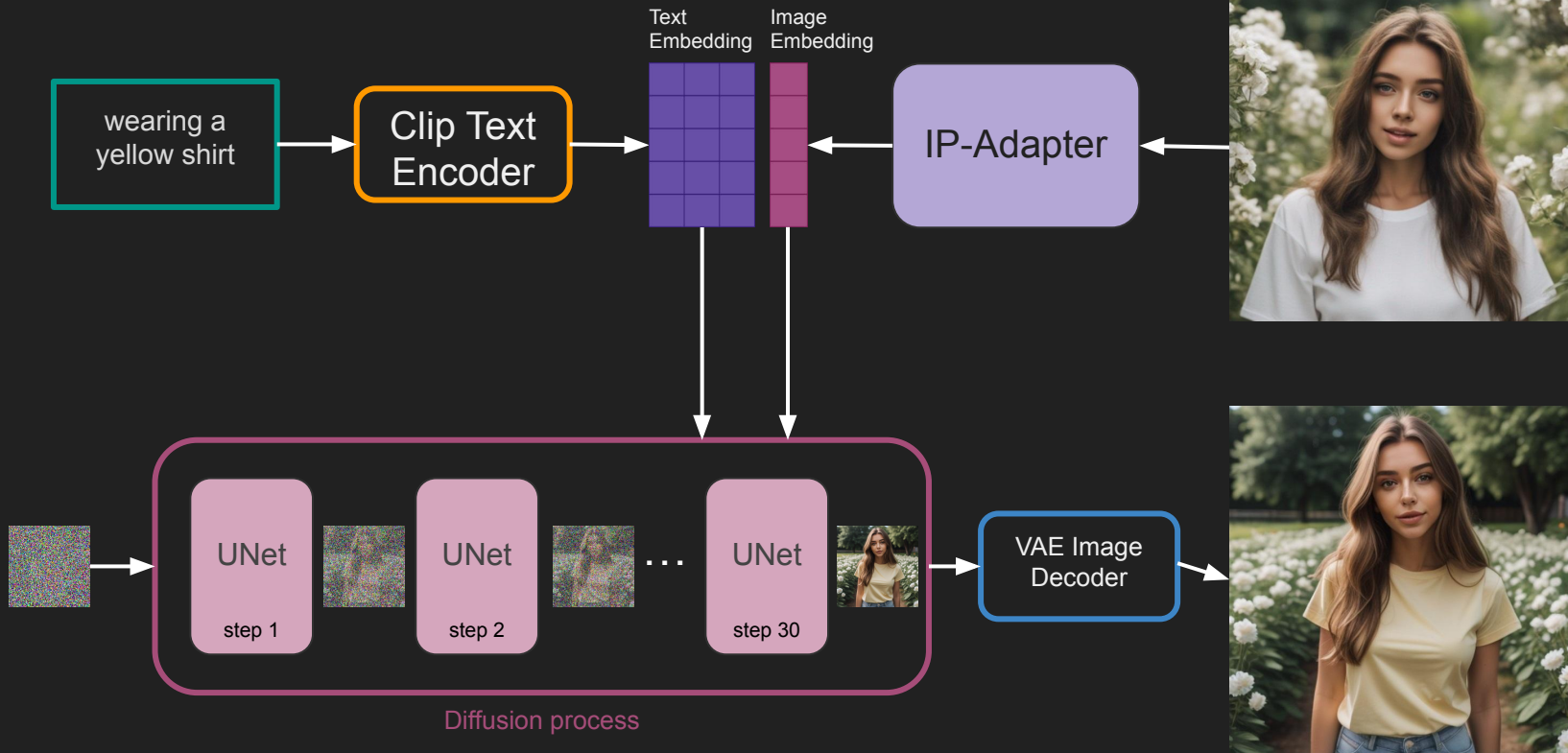
Segmentation



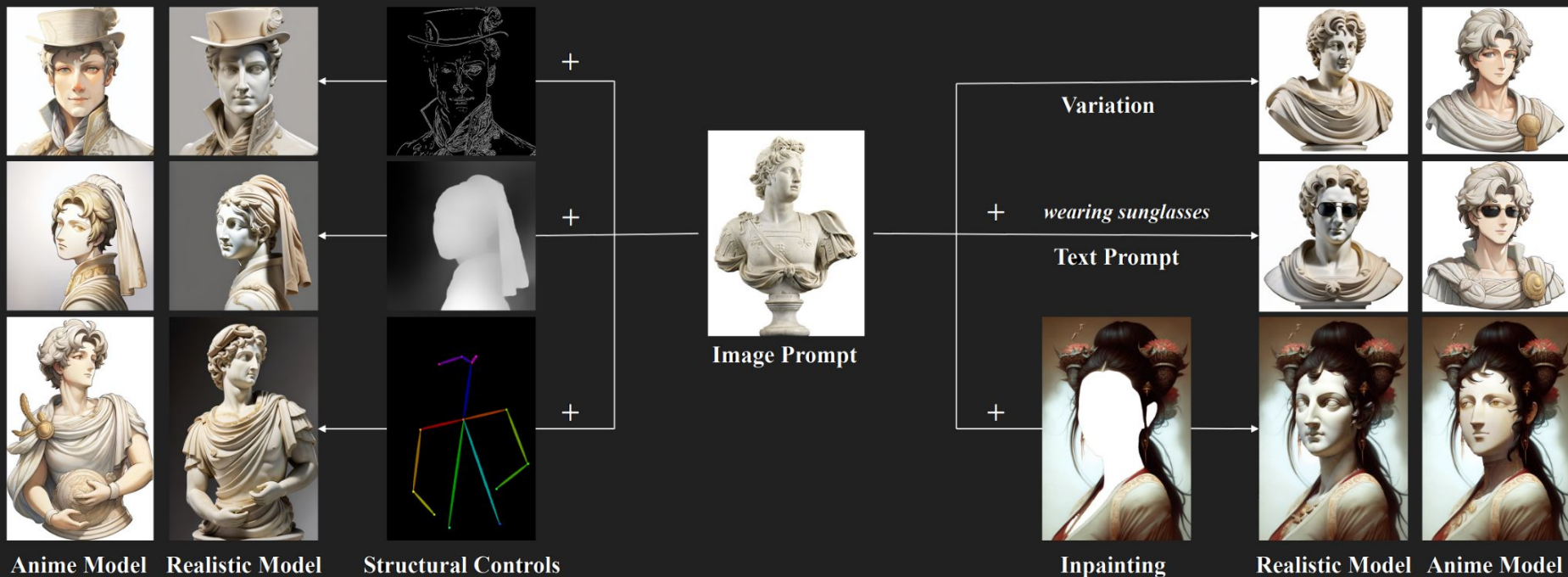
house

IP-Adapter

Image prompt



IP-Adapter



InstantID

style 1

style 2

style 3

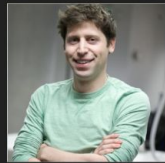
style 4

style 5

style 6

style 7

style 8



MultiDiffusion



“blurred image”

“a small bird”

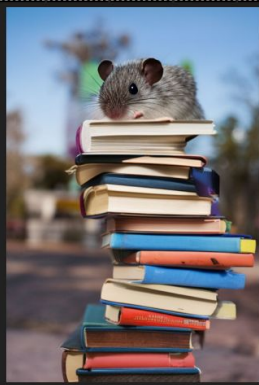
“a tree truck”



“blurred image”

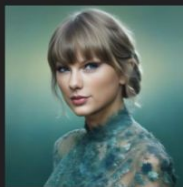
“a mouse”

“a pile of books”



Multi-concept Generation

LoRa



LoRa



a man and a woman on the street

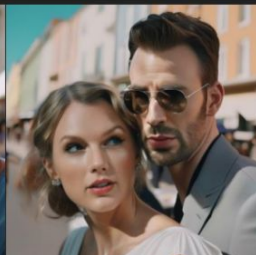
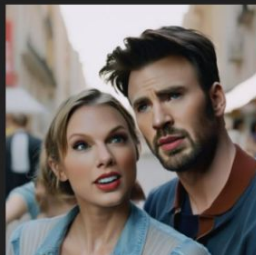
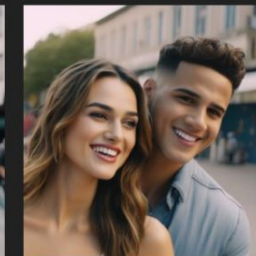
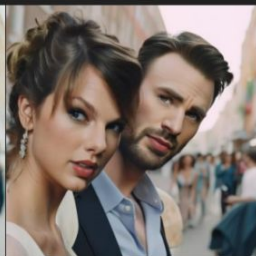
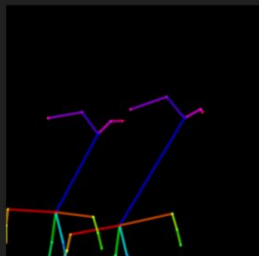
InstantID



InstantID



ControlNet
Human Pose



Спасибо за внимание!

Гайнетдинов А.Ф.

Team Lead, Senior Deep Learning Engineer, VicMan LLC

zirgan.ainur@gmail.com