

# TEXT-TO-IMAGE GENERATION

CS 319 / DS 231 Computer Vision

Fall 2023, American University of Armenia

Lilit Galstyan  
Anna Misakyan  
Aida Martirosyan

# Table of Contents

- Description of the Project
- Explanation and Solution
- Results of the Project

# Description of the Project

## Motivation

- Interests
- Fastly growing world of data / Generative models

## Challenges

- Research / Finding relevant information
- Utilizing the tools

## Choices / Decisions

# What We Used:

## Diffusion Models:

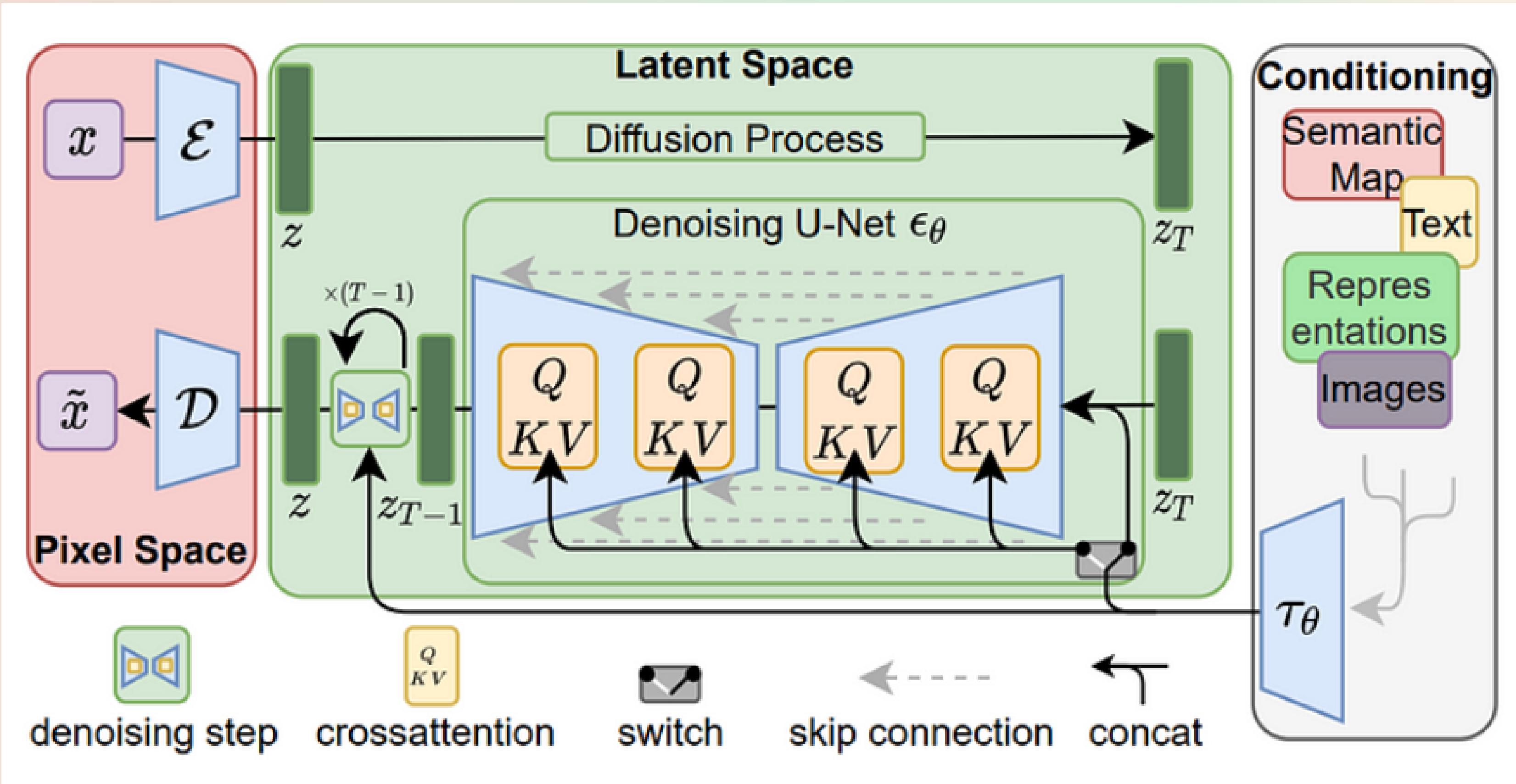
“diffusers” library

“Stable Diffusion” models

“DreamBooth” fine-tuning models

Rapunzel photo

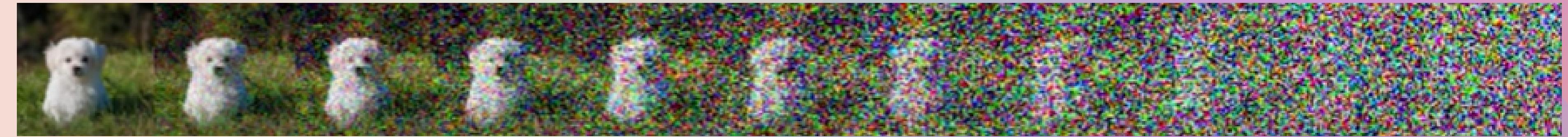
# Diffusion Model Architecture



# Linear Scheduler



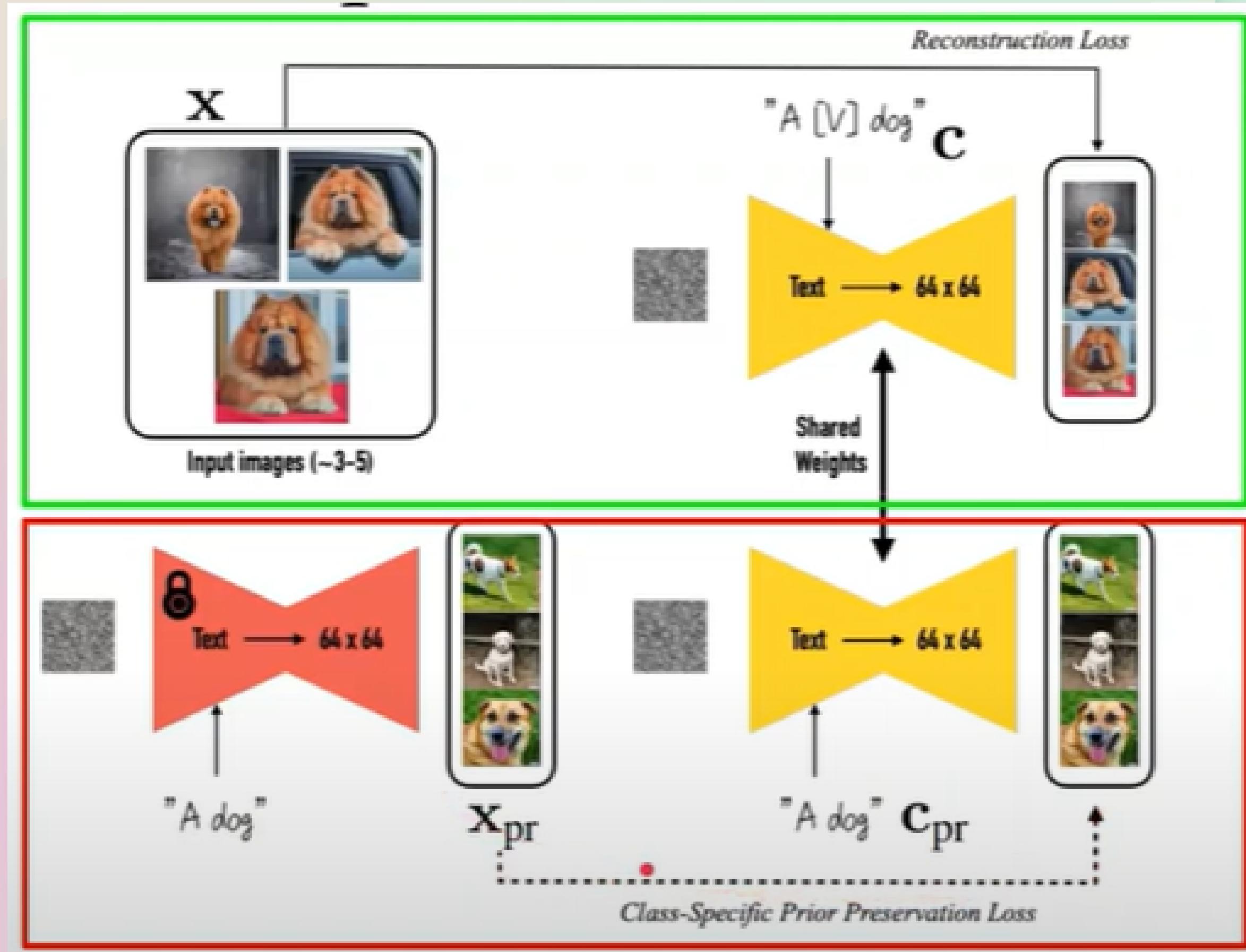
# Non - Linear Scheduler



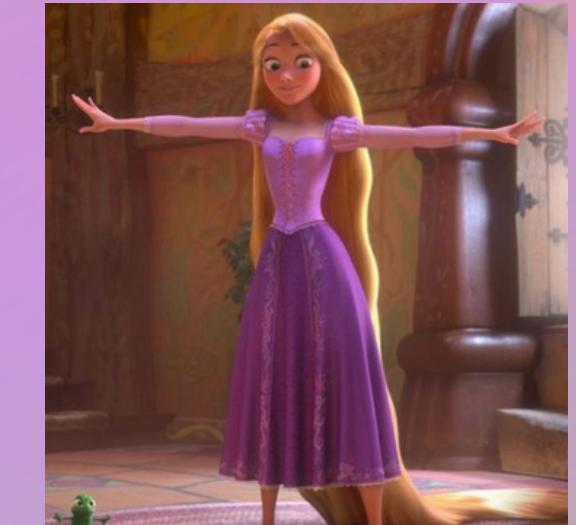
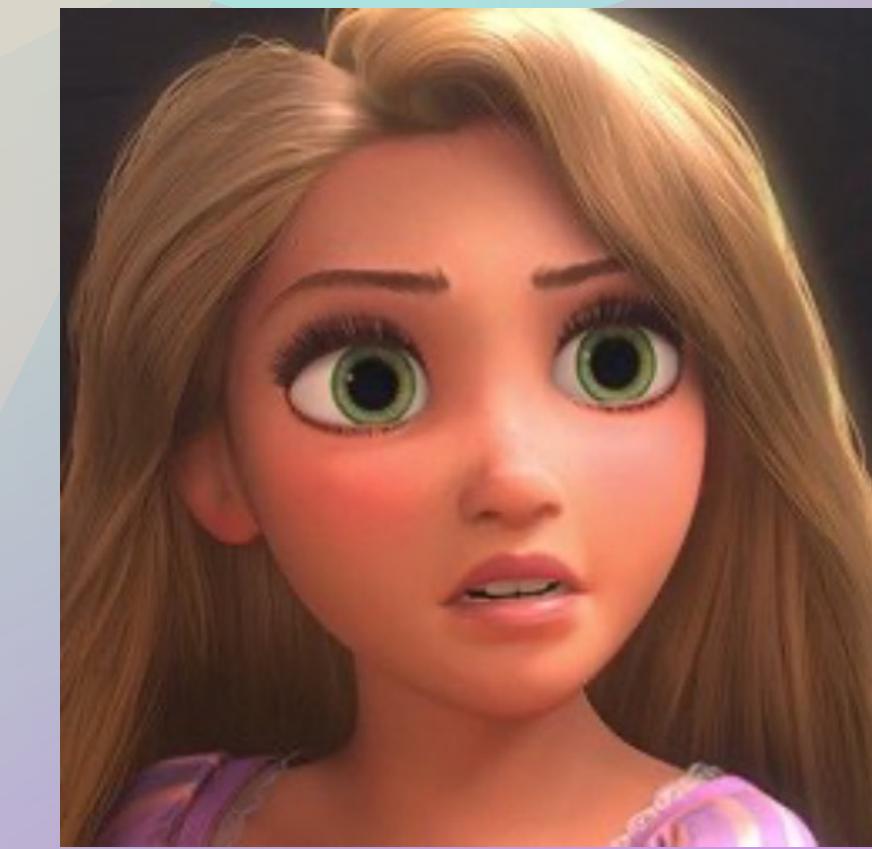
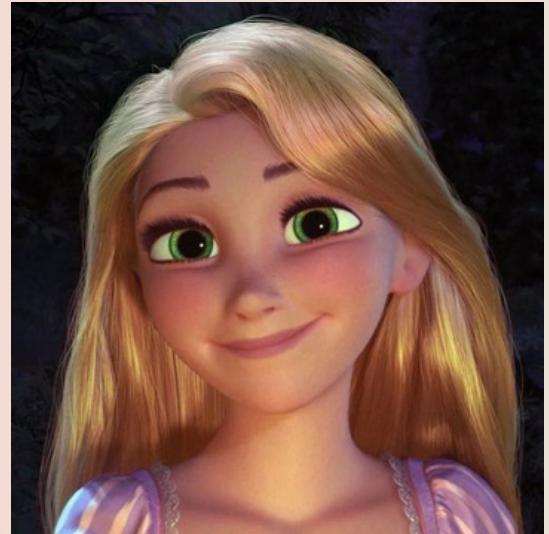
# What is DreamBooth?



# DreamBooth Architecture

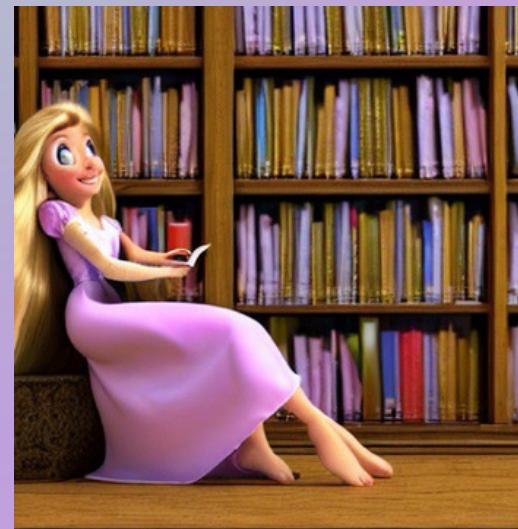
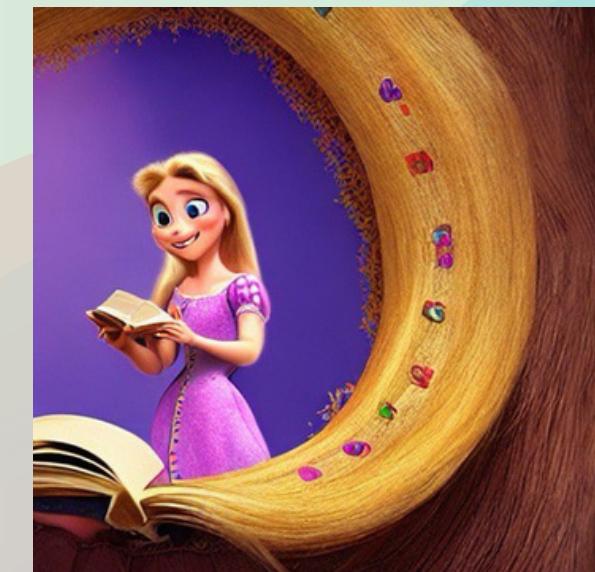


# Results - Original Images

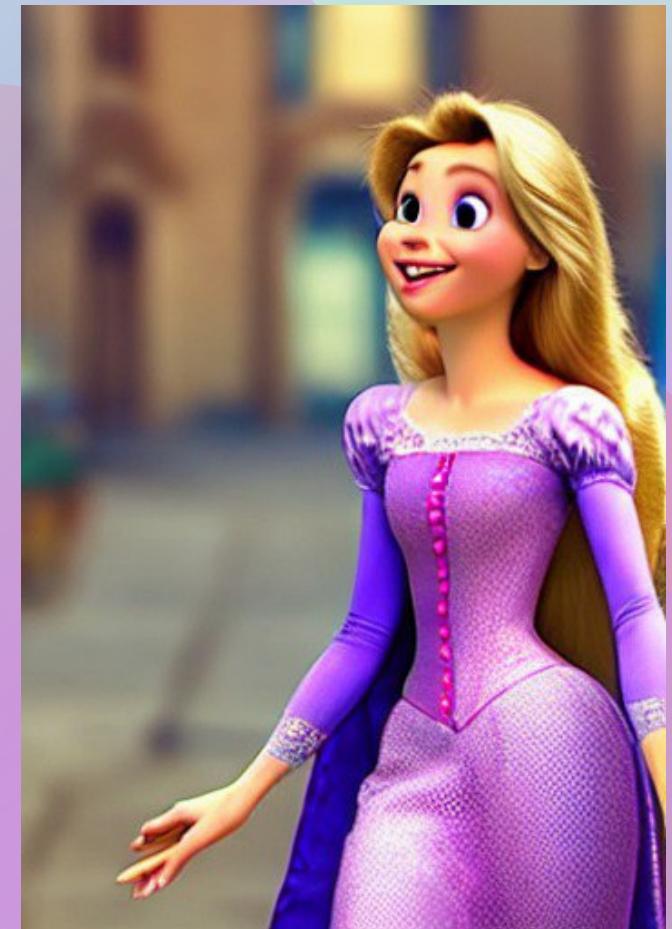


# Initial Results

- Unpreserved dress style

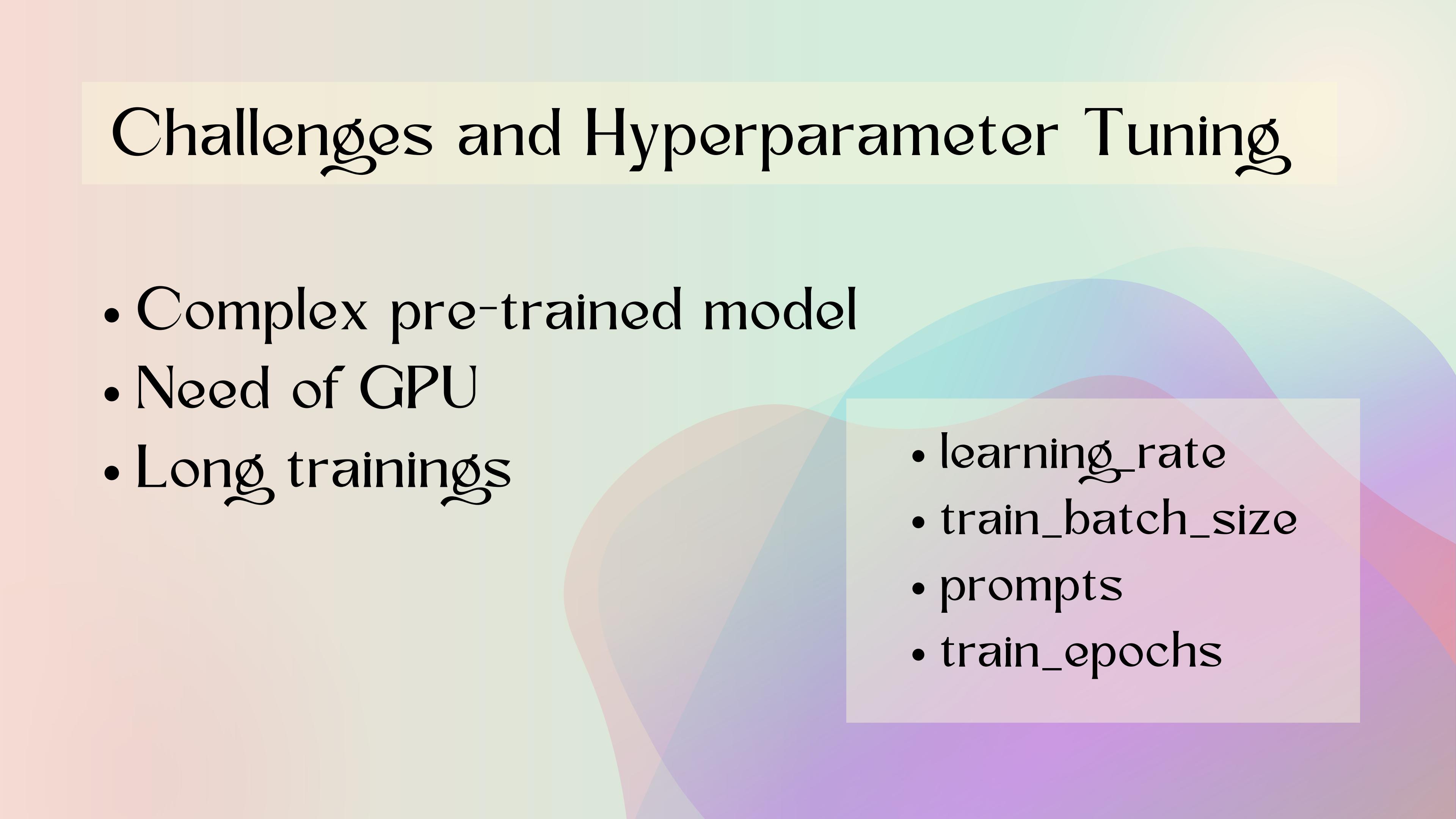


- Unpreserved body parts



# Challenges and Hyperparameter Tuning

- Complex pre-trained model
- Need of GPU
- Long trainings

- 
- learning\_rate
  - train\_batch\_size
  - prompts
  - train\_epochs

# Final Results 1

- driving a car



- reading a book



# Final Results 2

- preparing for Computer Vision course with computer
- looking in the mirror
- speaking on the phone



# THANK YOU!

