

موضوع

Gen CV

Generative Computer Vision

جيCV - موضوع

Prompt → Image!

Diffusion!

Generated Image



Image!

تولید تصویر از متن (Text-to-Image Generation)

تولید تصاویر واقع‌گرایانه یا هنری از توصیف‌های متنی

Stable Diffusion / DALL·E 2 / Imagen

Dane?

Prompt

A vibrant underwater scene with a playful dolphin swimming through a coral reef |

Generate

ترمیم تصویر (Image Inpainting)

بازسازی قسمت‌های حذف شده، پنهان شده یا خراب شده از یک تصویر

Two-Stage GAN

~~Two-Stage GAN~~

prompt
پریم

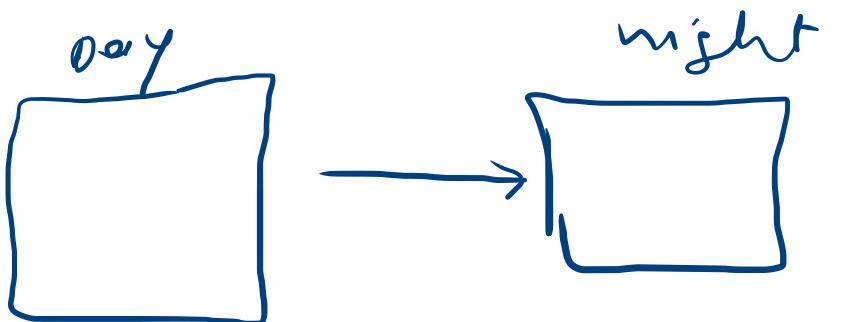


(a) Input

(b) Ground Truth

(c) RFR

(d) MFR-GAN



تبدیل سبک تصویر (Style Transfer)

انتقال سبک هنری یک تصویر به تصویر دیگر

CycleGAN



1 Upload photo

The first picture defines the scene you would like to have painted.



2 Choose style

Choose among predefined styles or upload your own style image.



3 Submit

Our servers paint the image for you. You get an email when it's done.



original



style



افزایش وضوح تصویر (Super-Resolution)

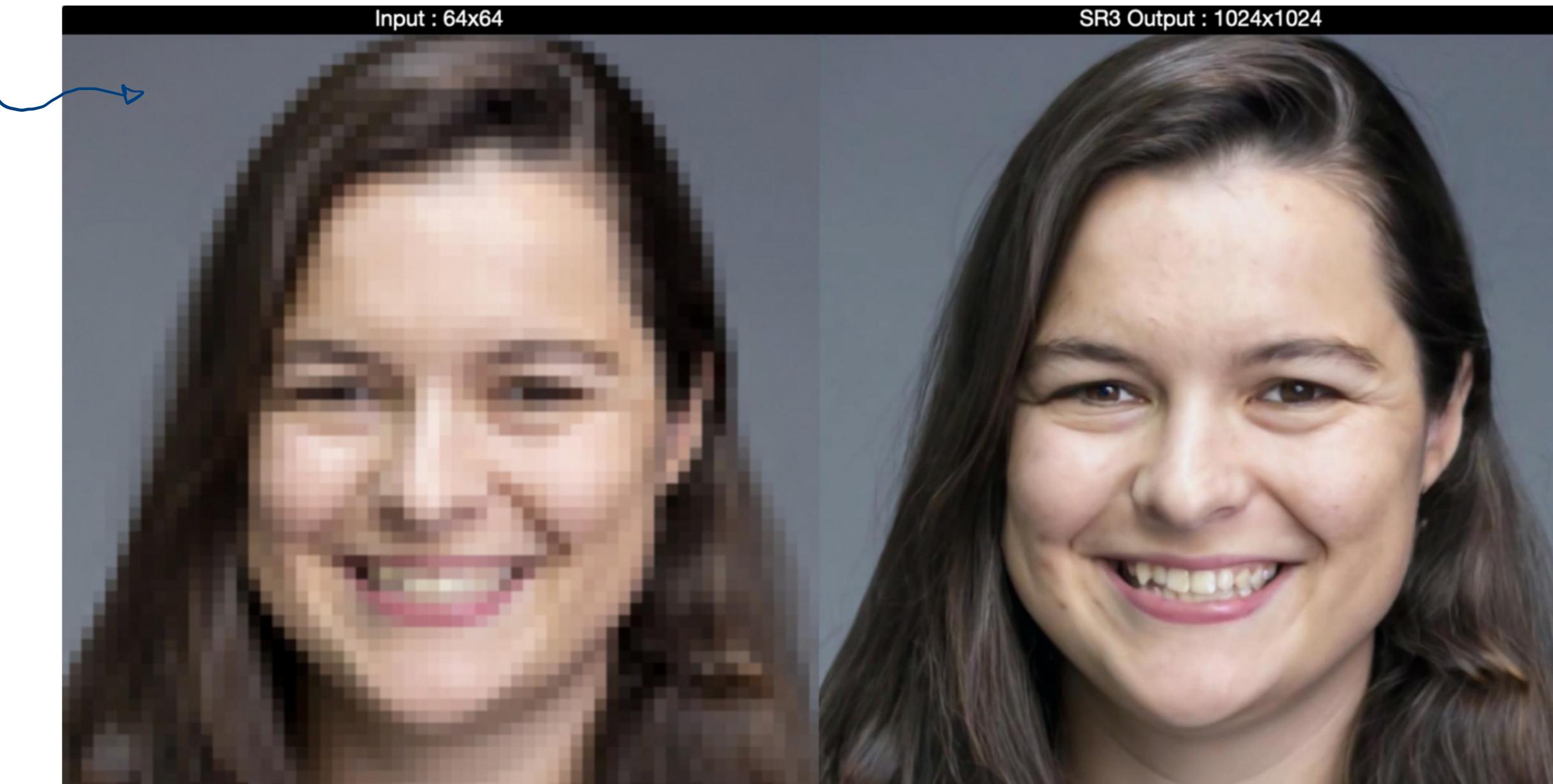
افزایش کیفیت یا ابعاد تصاویر بدون از دست دادن جزئیات

SRGAN

input

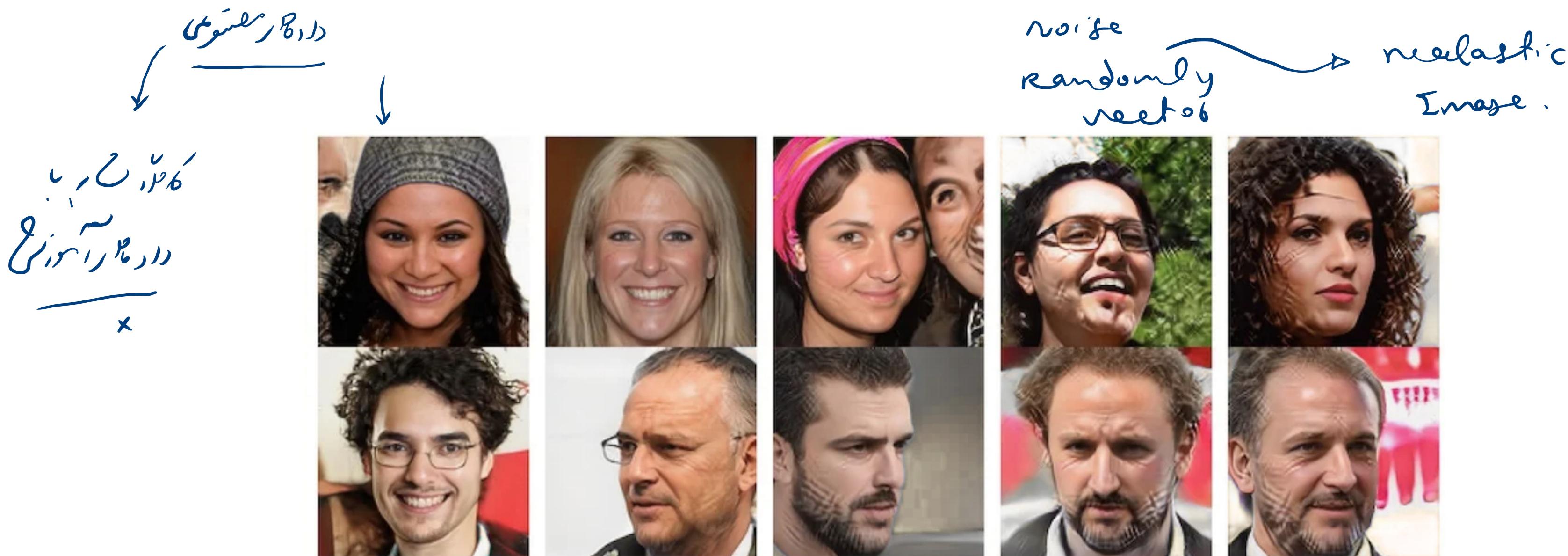
SR

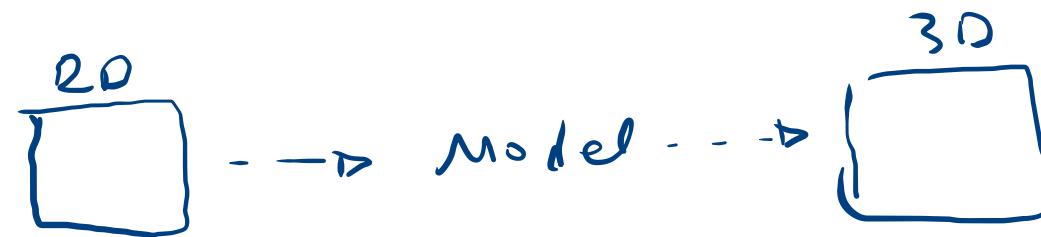
output



تولید داده مصنوعی (Data Augmentation / Synthetic Data Generation)

تولید تصاویر مصنوعی برای افزایش حجم داده آموزشی *

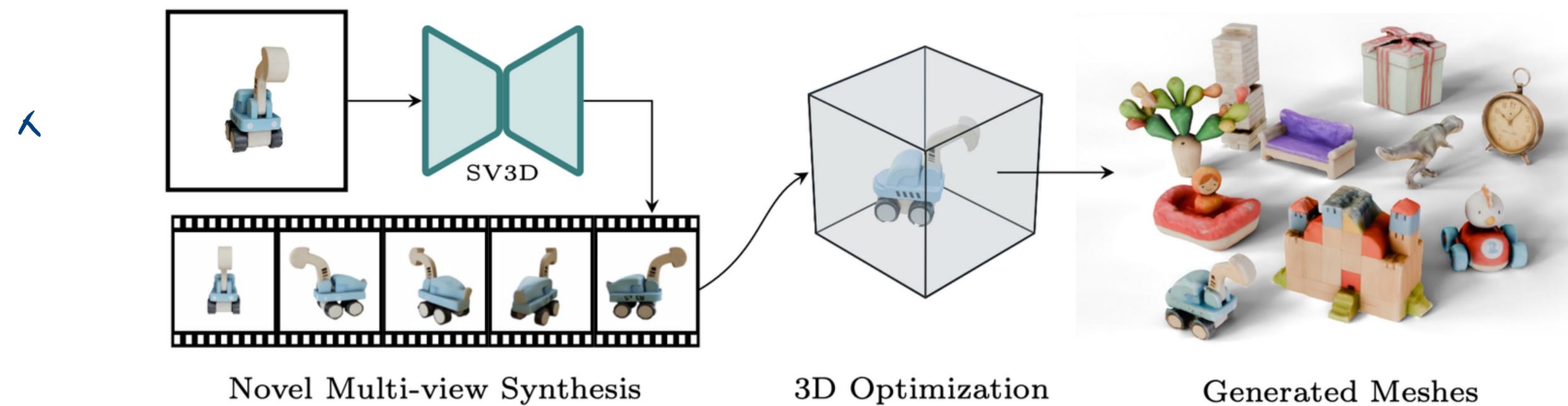




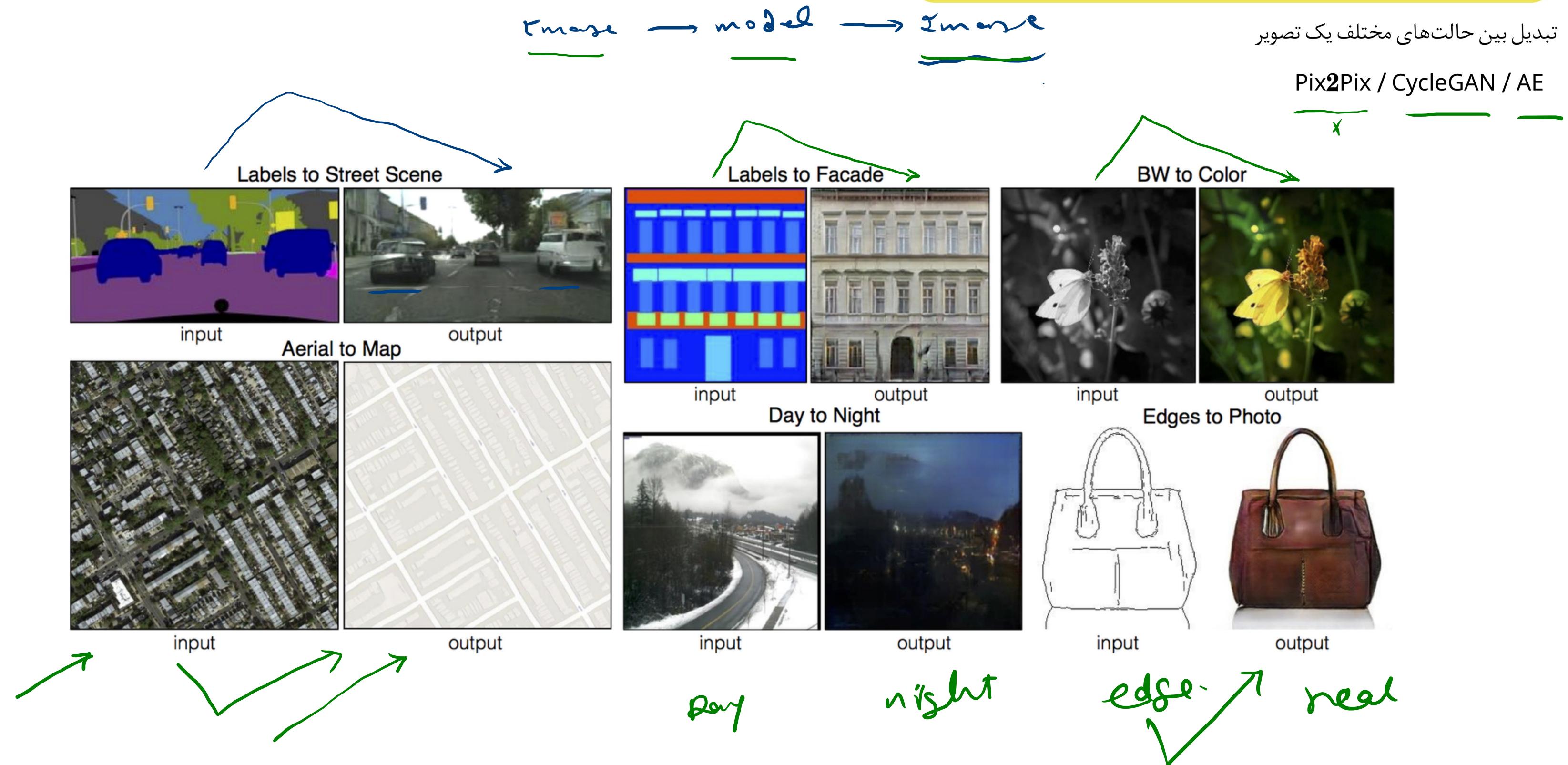
بازسازی ۳ بعدی از تصاویر دو بعدی (3D Reconstruction / Novel View Synthesis)

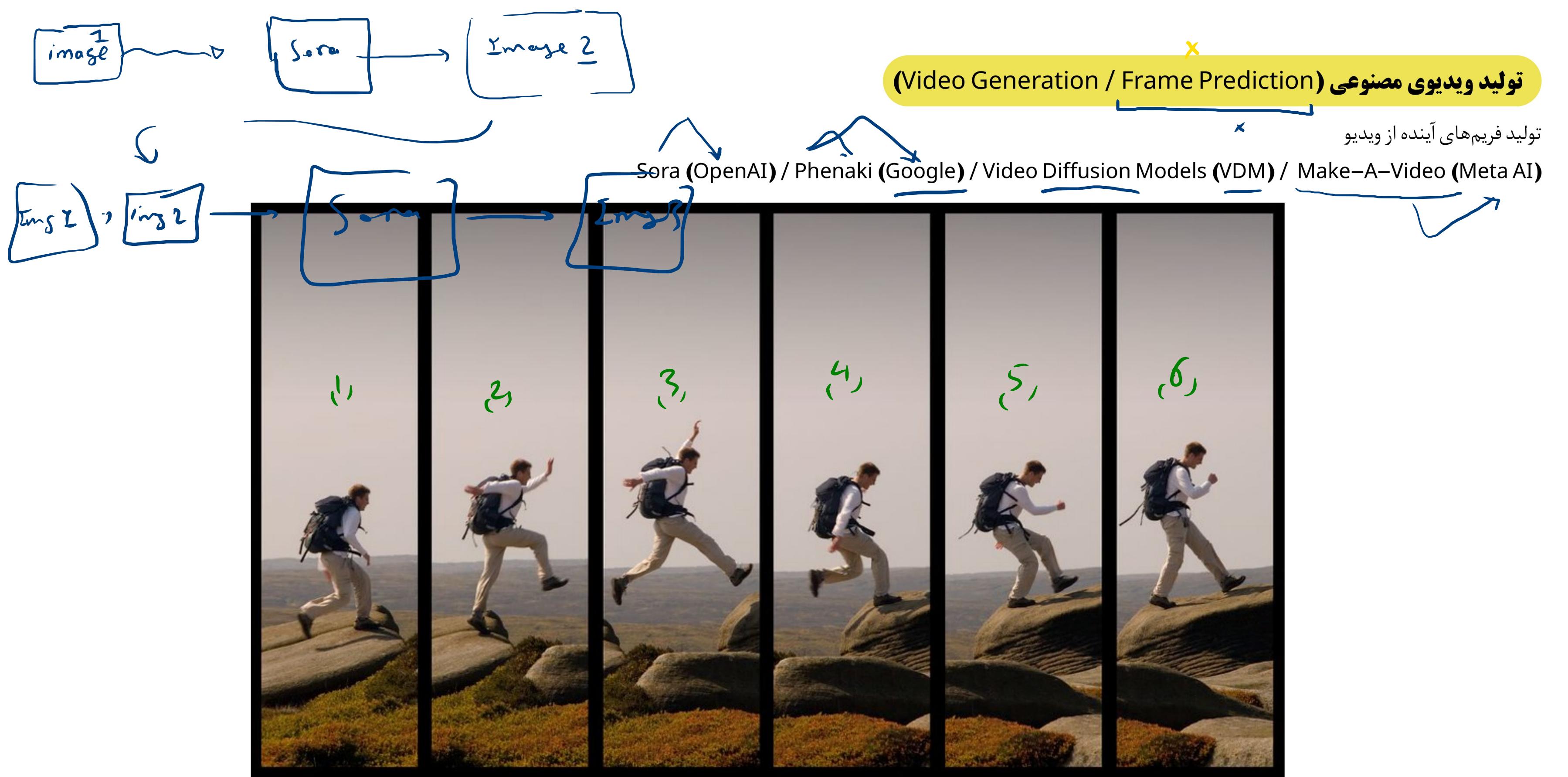
بازسازی صحنه یا شئ سه بعدی از تصاویر مختلف یا از یک تصویر

NeRF / Instant-NGP



ترجمه تصویر به تصویر (Image-to-Image Translation)

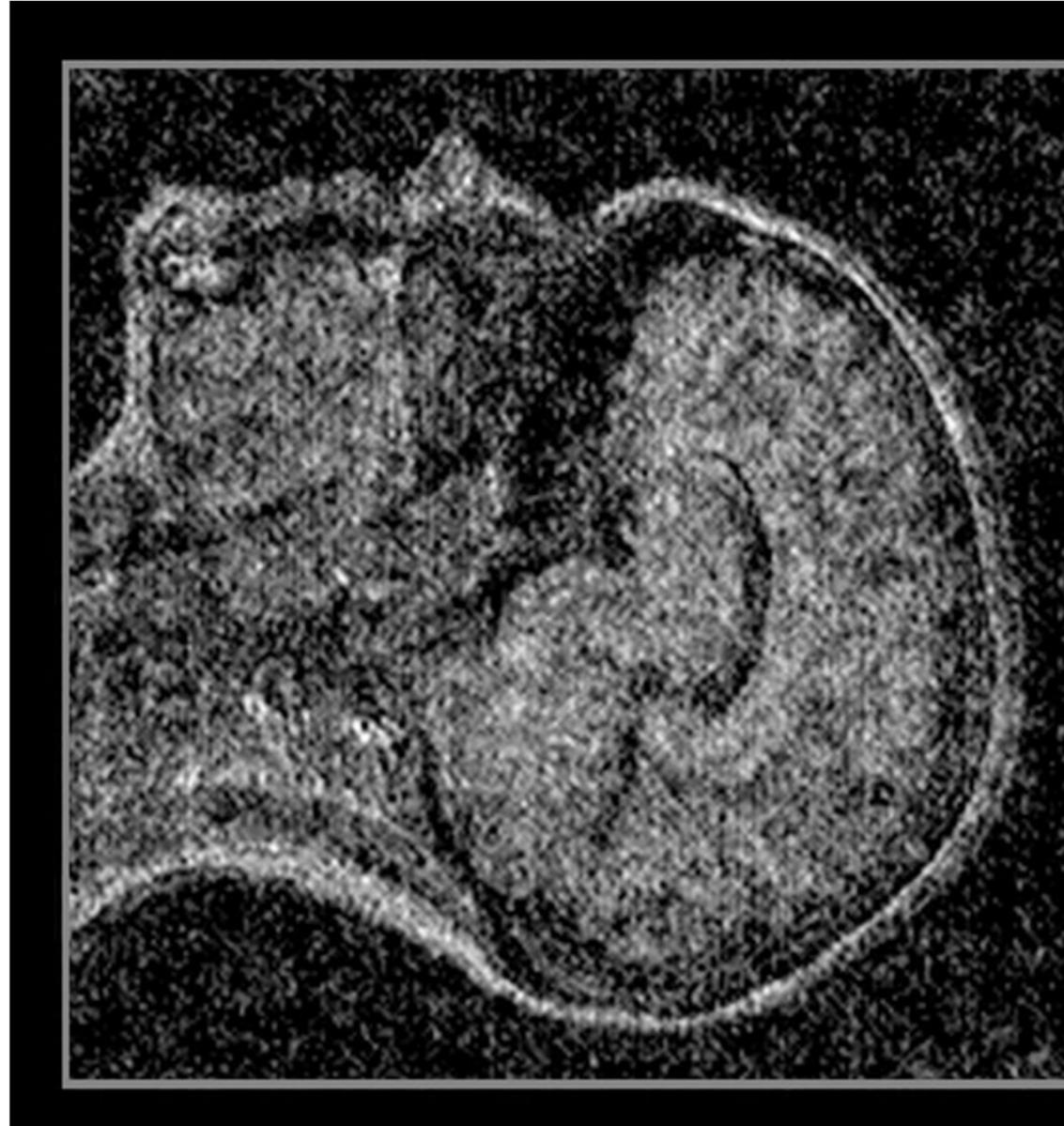




α

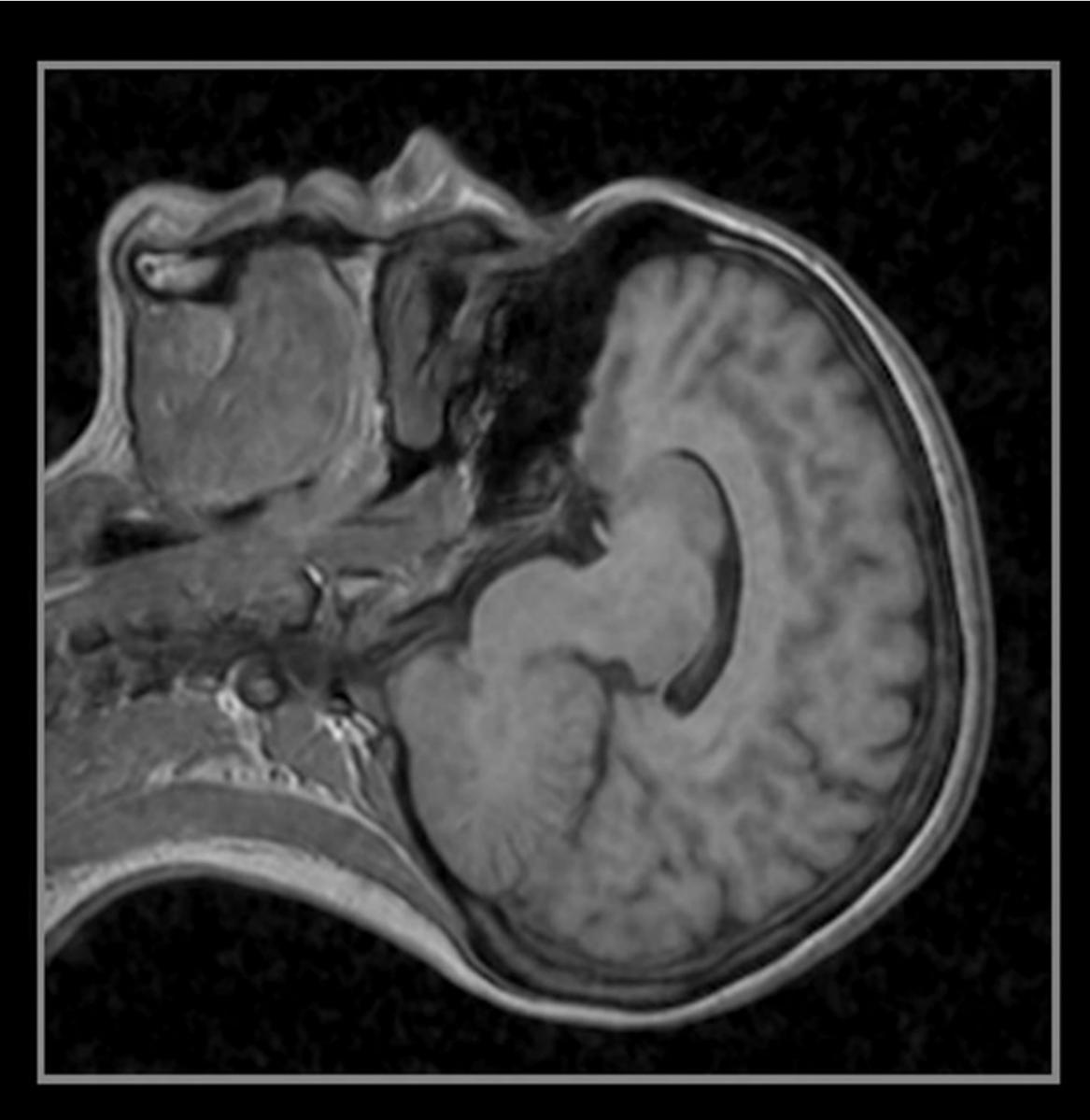
بازیابی تصویر از نویز (Denoising)

input



Sora (OpenAI) / Phenaki (Google) / Video Diffusion Models (VDM) / Make-A-Video (Meta AI)

output



بازیابی تصویر از نویز (Denoising)

x

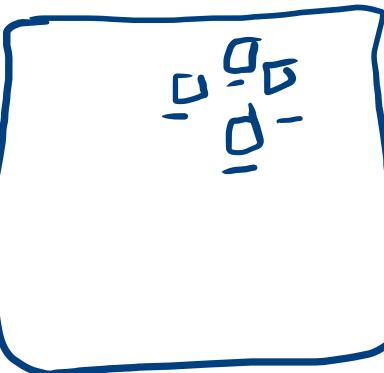
تولید فریم‌های آینده از ویدیو

VAE
x

۱۰. رفزنه لدی درین کار

watermark ✓

6

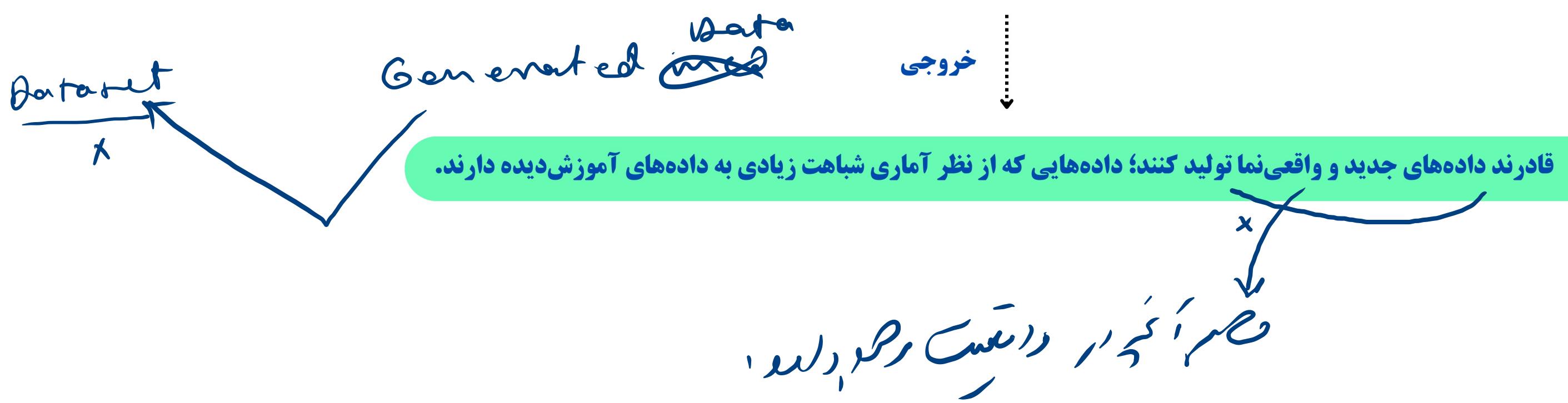
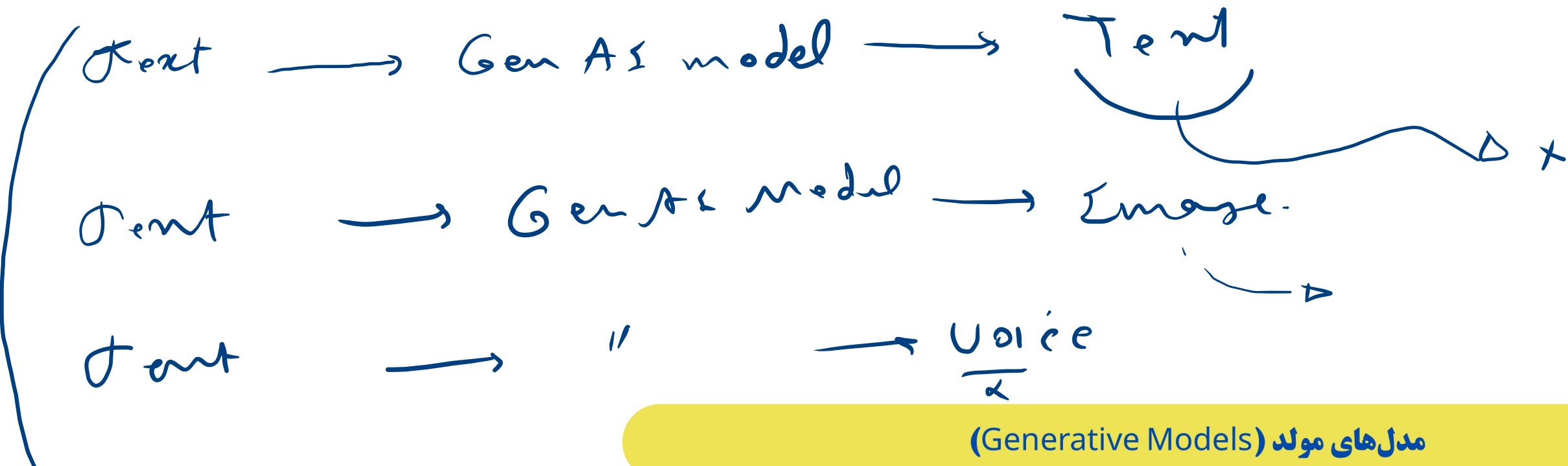


A red arrow points from left to right, overlaid on a black and white digital noise background. The arrow points to a red circular area containing binary code.

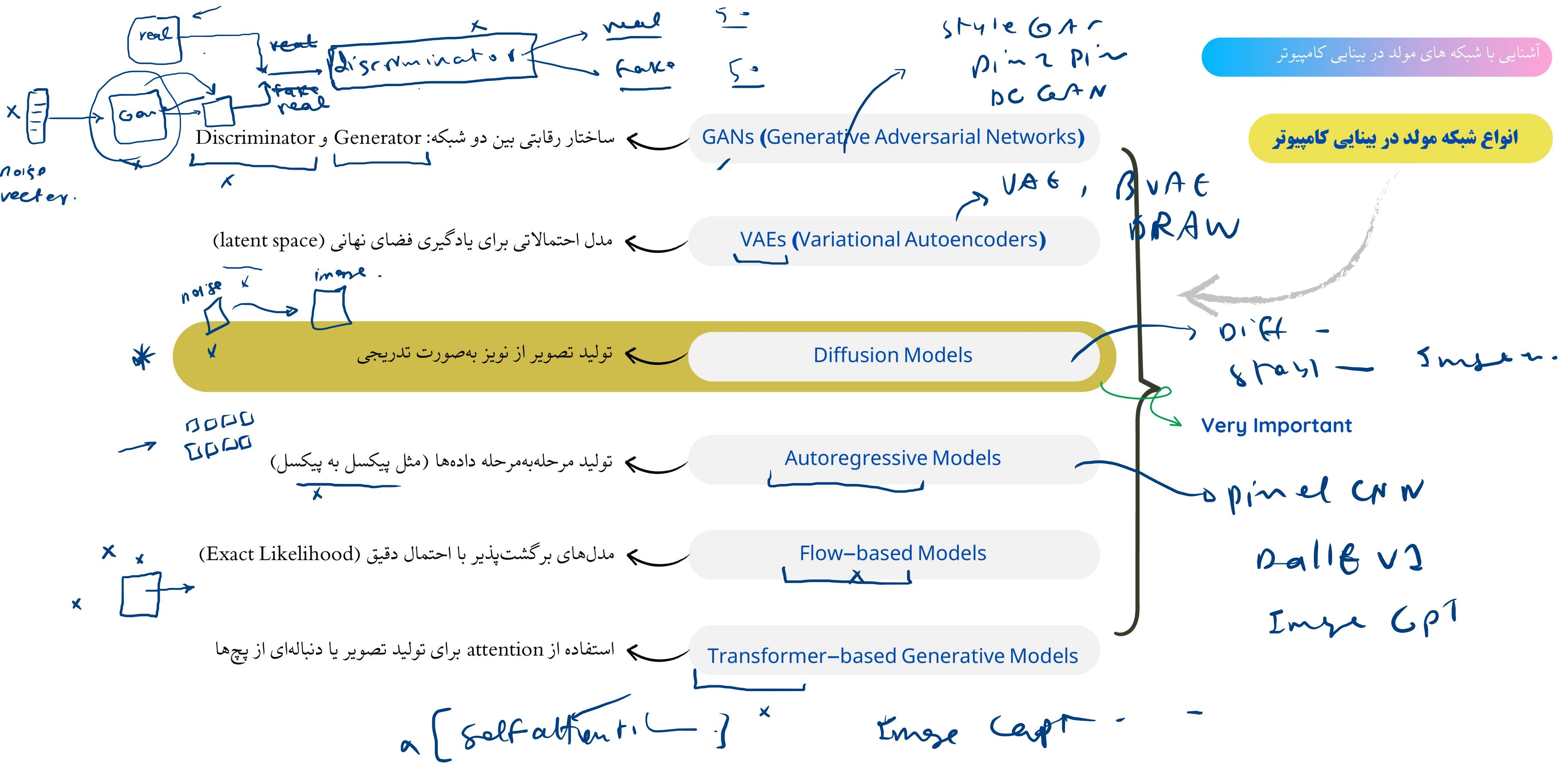
پنهان نگاری و رمزگذاری تصویری (Steganography & Watermarking)

پنهان کردن داده (مثل متن یا تصویر) درون یک تصویر به کمک شبکه‌های مولد

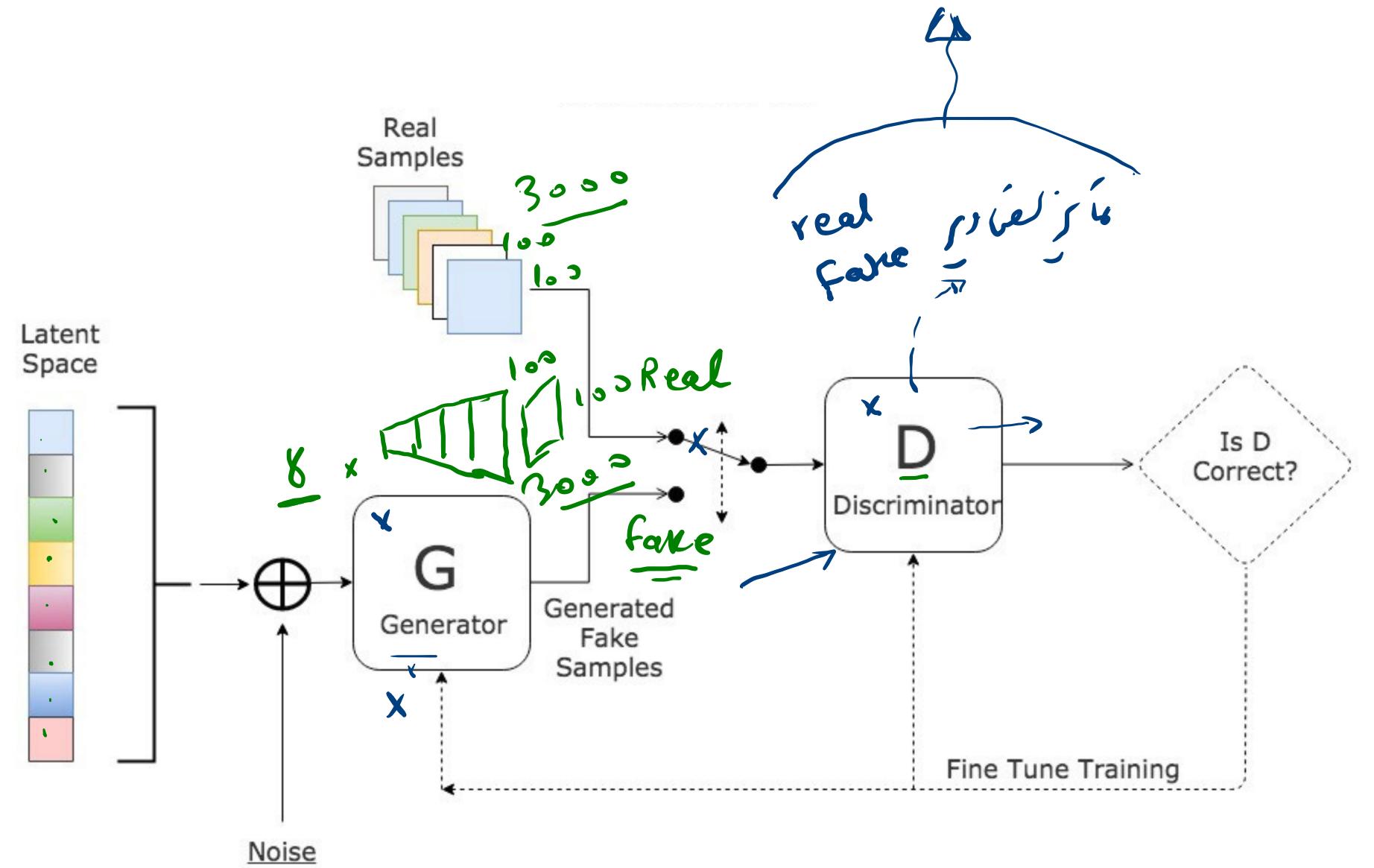
SteganoGAN (2020) / DeepStamp



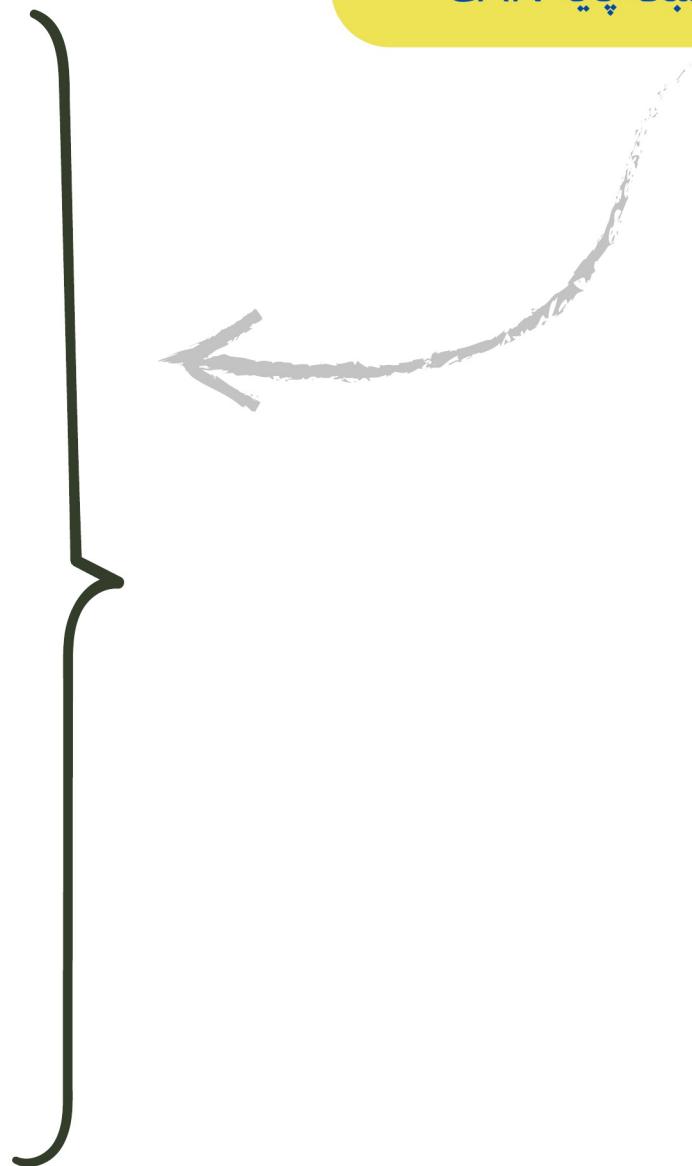
انواع شبکه مولد در بینایی کامپیوتر

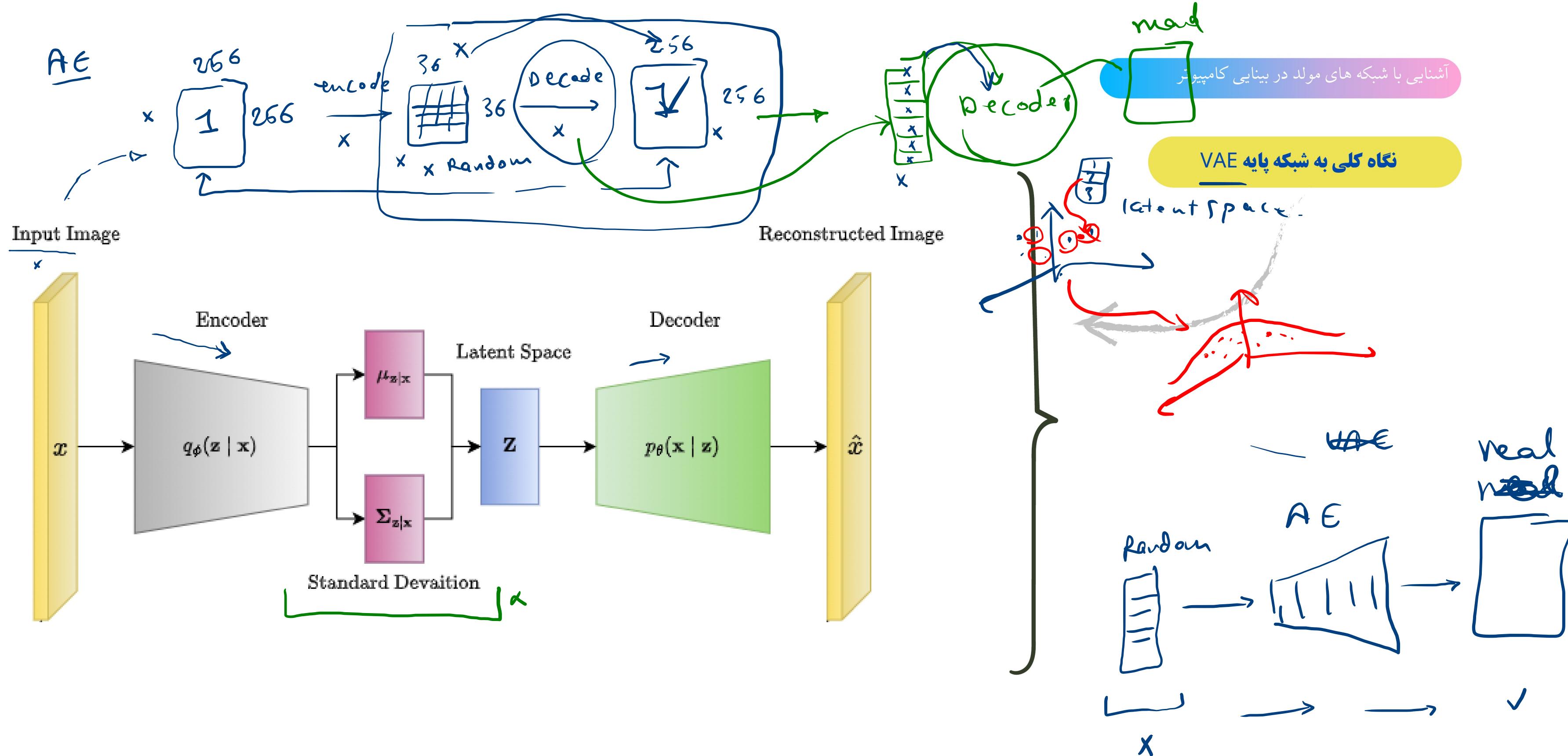


GAN گان



نگاه کلی به شبکه پایه GAN

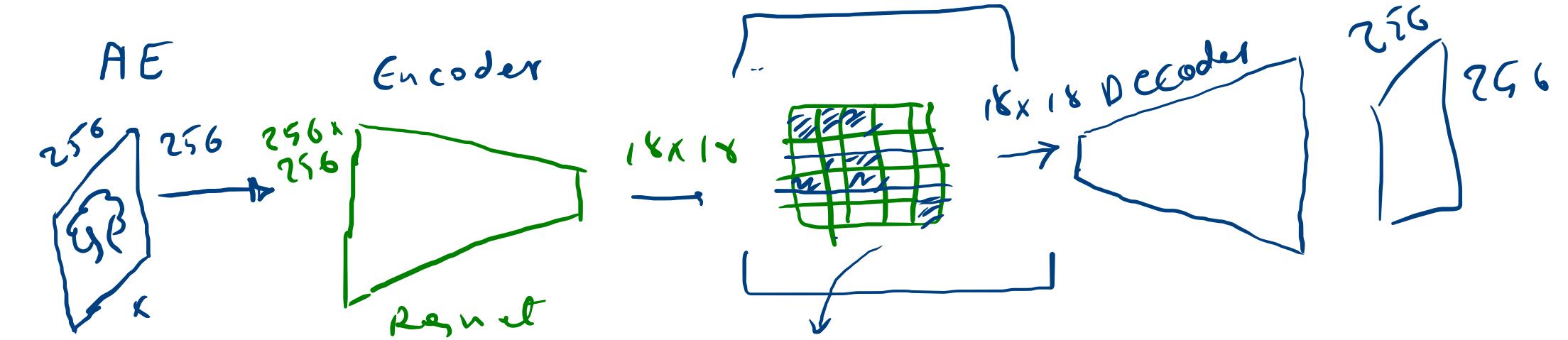




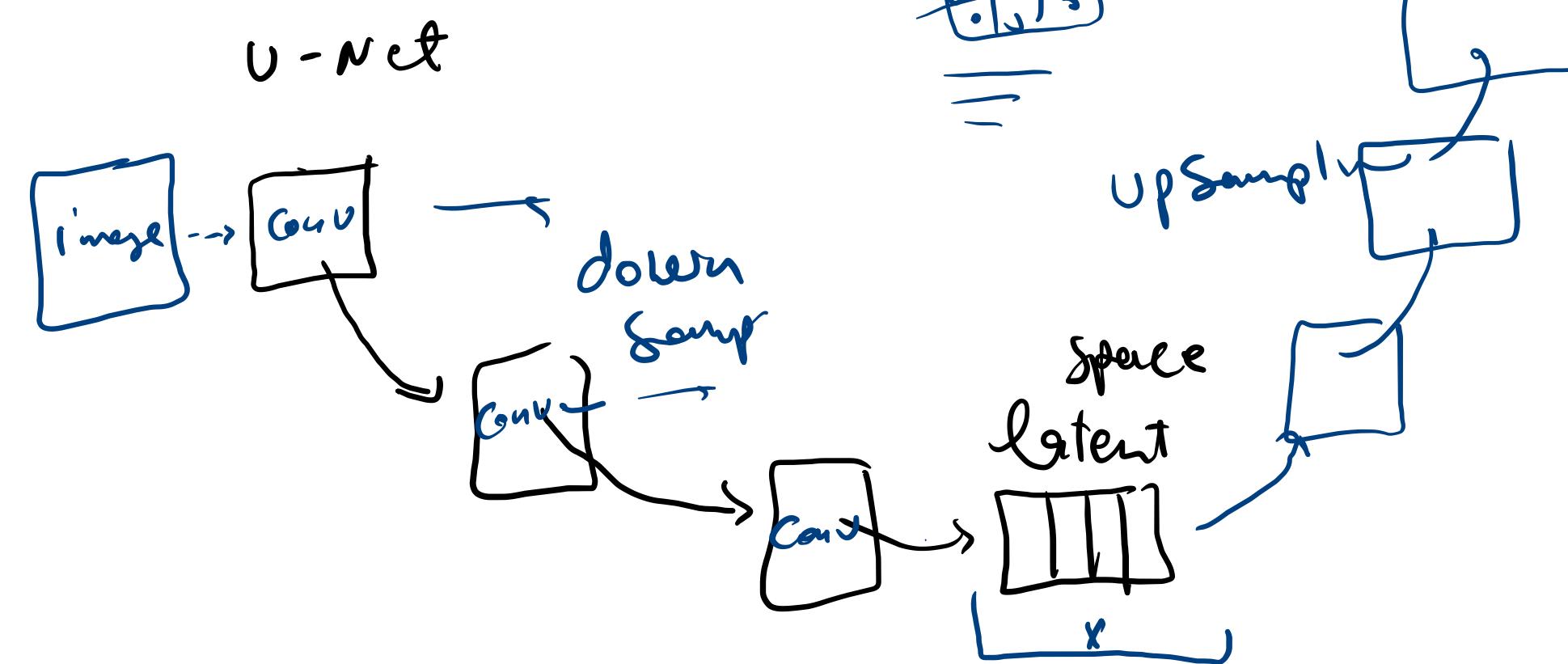
Latent Space.

فرموده پیش از معرفی

[!سازی]



U-Net



VAE \times → Face Image Generation → انتزاع

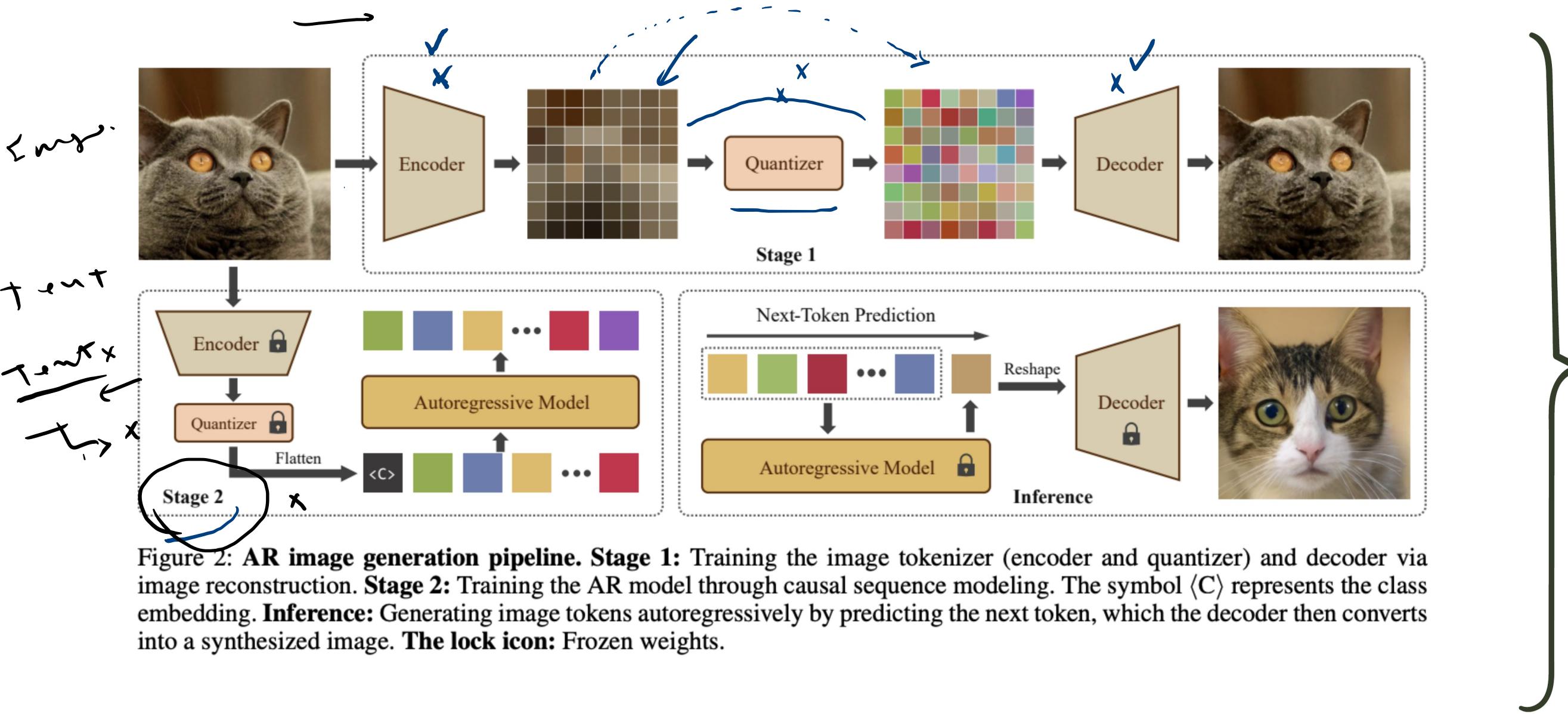
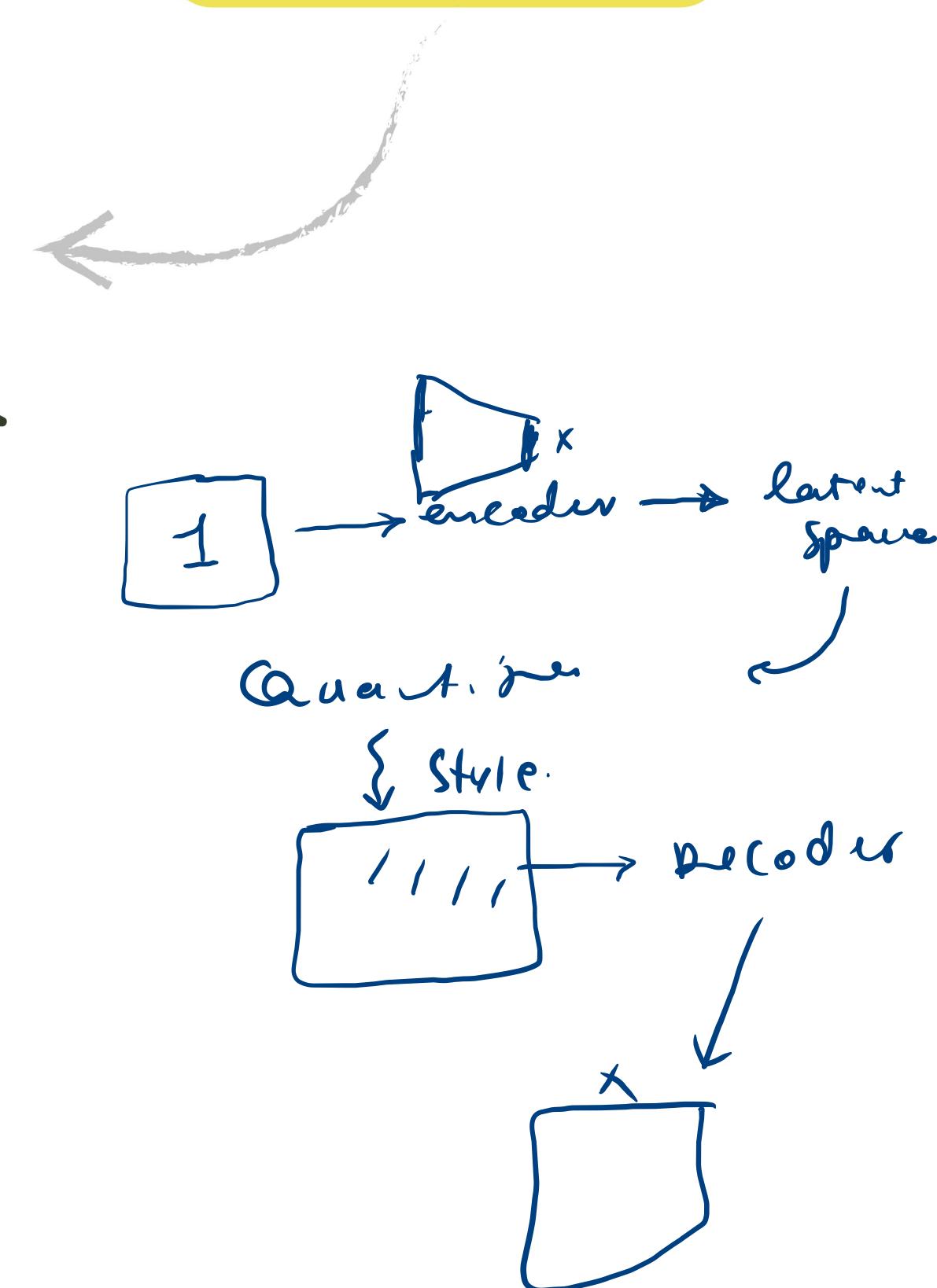
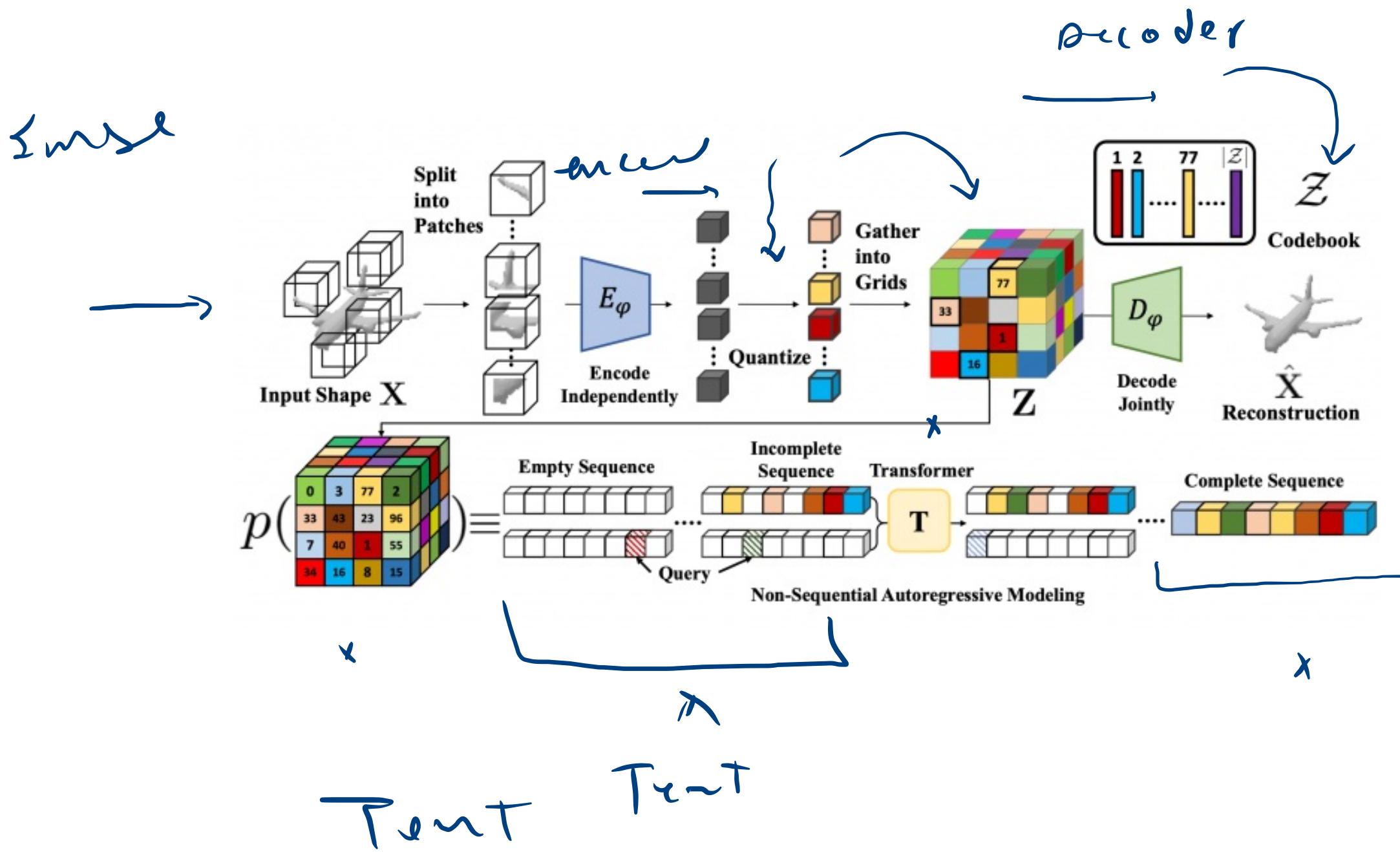


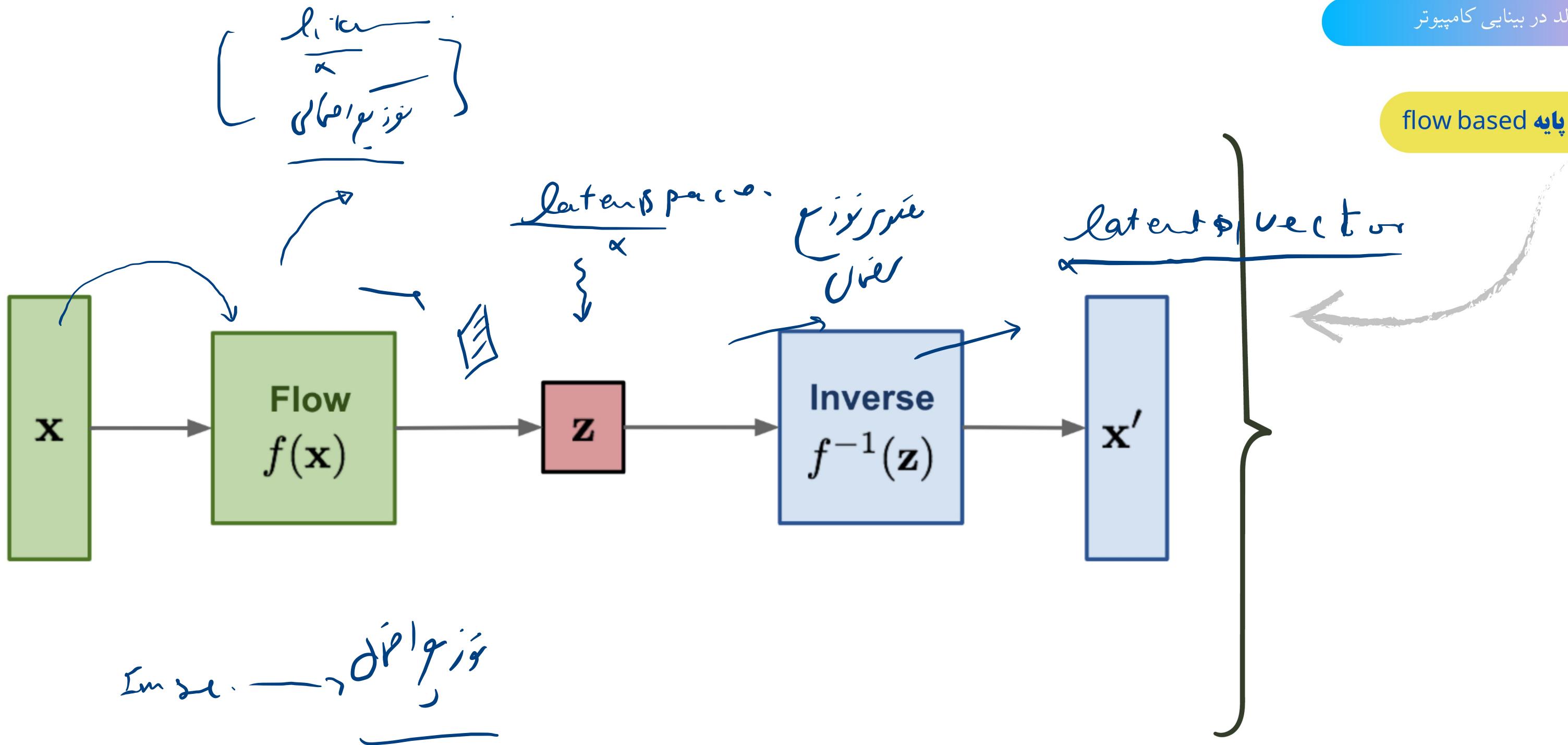
Figure 2: AR image generation pipeline. **Stage 1:** Training the image tokenizer (encoder and quantizer) and decoder via image reconstruction. **Stage 2:** Training the AR model through causal sequence modeling. The symbol $\langle C \rangle$ represents the class embedding. **Inference:** Generating image tokens autoregressively by predicting the next token, which the decoder then converts into a synthesized image. **The lock icon:** Frozen weights.

نگاه کلی به شبکه پایه

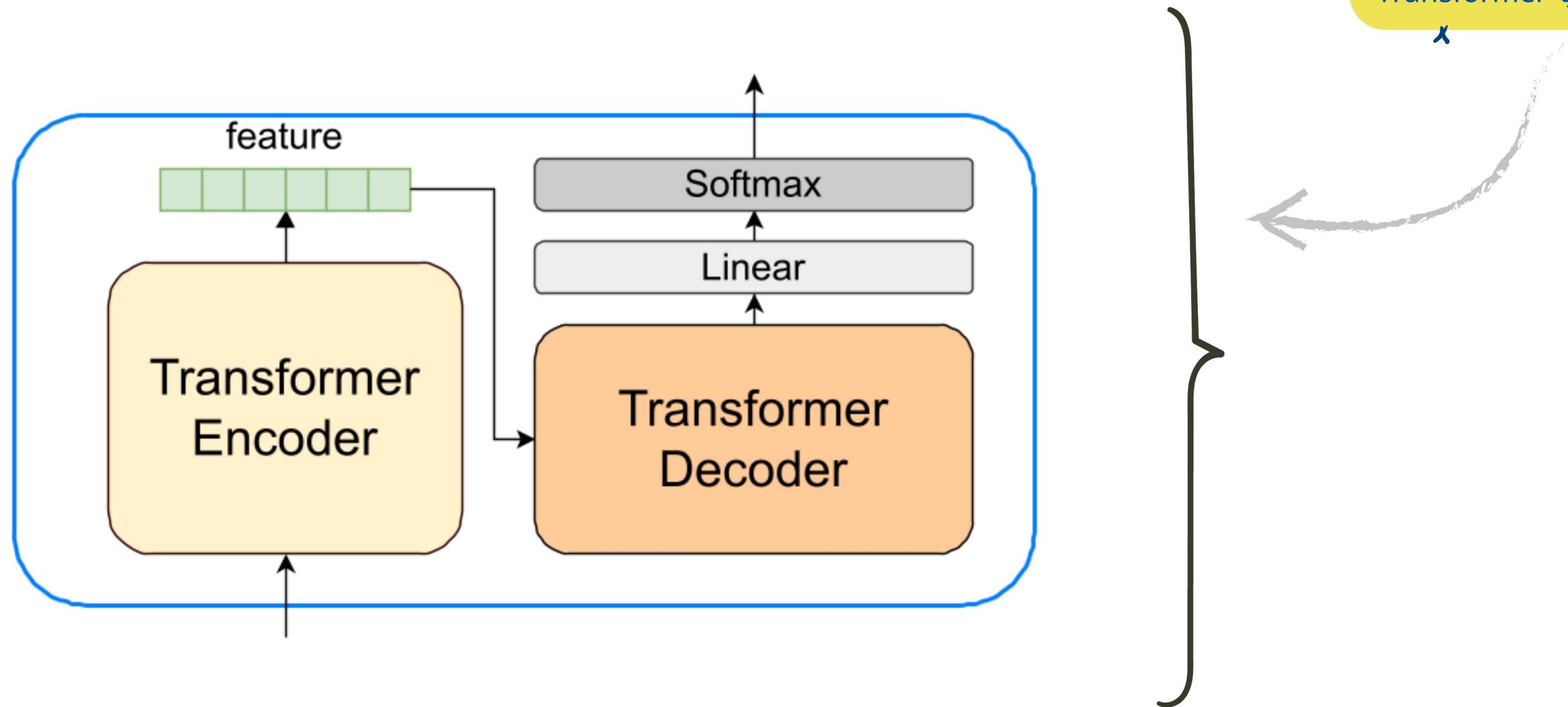




نگاه کلی به شبکه پایه



نگاه کلی به شبکه پایه Transformer



موضوع

Diffusion

جلسه بعد