

# Deep Generative Models

## Lecture 14

Roman Isachenko

Moscow Institute of Physics and Technology  
Yandex School of Data Analysis

2024, Autumn

# Outline

## 1. Latent models

Score-based models

Autoregressive models

## 2. The worst course overview

# Outline

## 1. Latent models

Score-based models

Autoregressive models

## 2. The worst course overview

# Outline

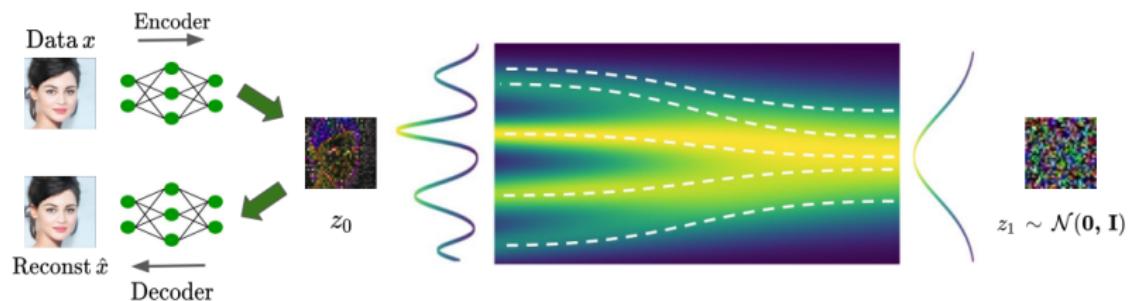
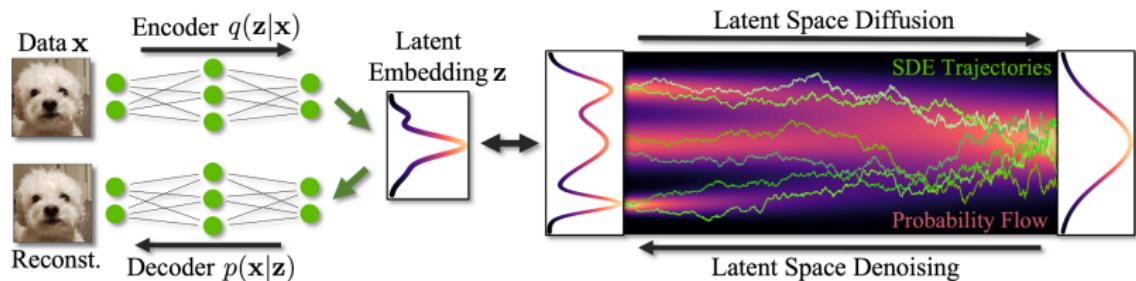
## 1. Latent models

Score-based models

Autoregressive models

## 2. The worst course overview

# Latent models



Dao Q. et al. *Flow Matching in Latent Space*, 2023

NeurIPS 2023 Tutorial: Latent Diffusion Models: Is the Generative AI Revolution Happening in Latent Space?

# Outline

## 1. Latent models

Score-based models

Autoregressive models

## 2. The worst course overview

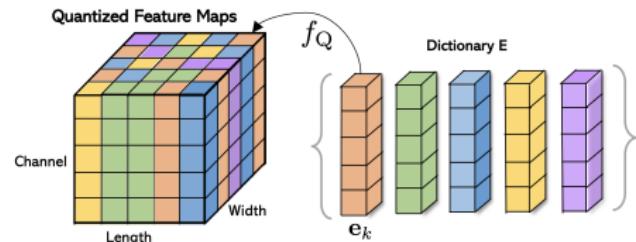
# Vector Quantized VAE (VQ-VAE)

## Vector quantization

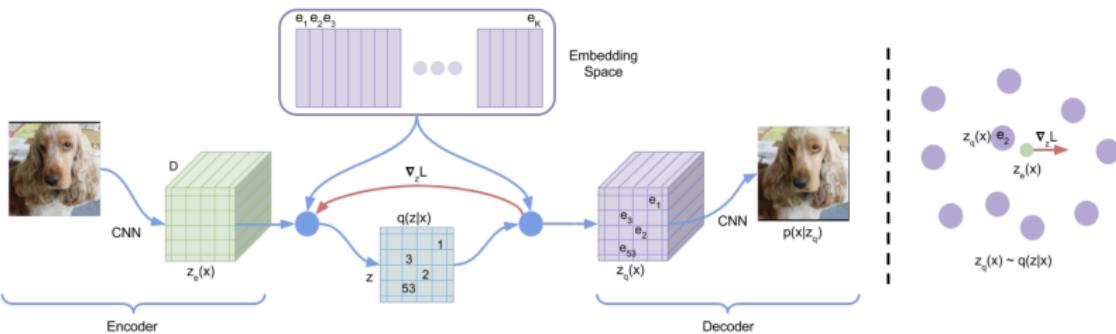
Define the dictionary space  $\{\mathbf{e}_k\}_{k=1}^K$ , where  $\mathbf{e}_k \in \mathbb{R}^C$ ,  $K$  is the size of the dictionary.

$$\mathbf{z}_q = \mathbf{q}(\mathbf{z}) = \mathbf{e}_{k^*}$$

$$\text{Here } k^* = \arg \min_k \|\mathbf{z} - \mathbf{e}_k\|.$$



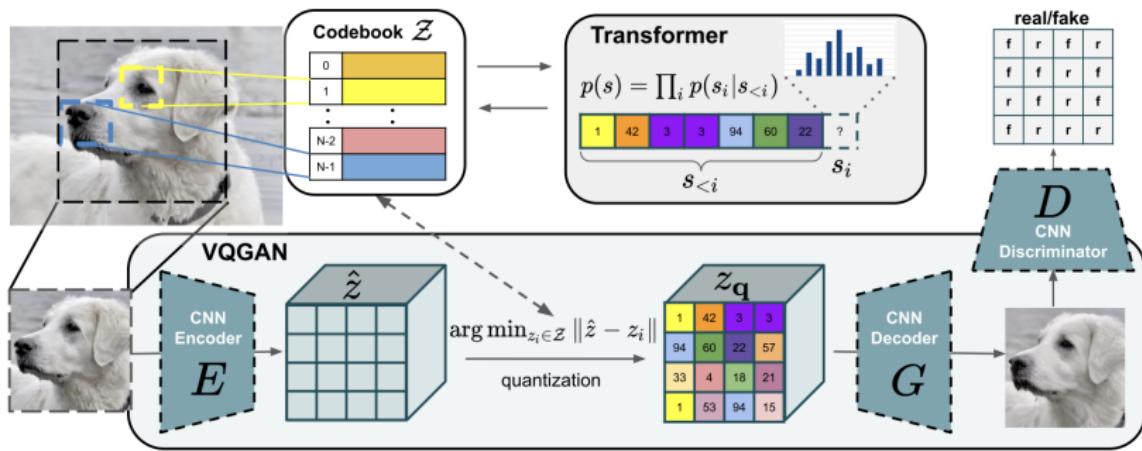
$$\mathcal{L}_{\phi, \theta}(\mathbf{x}) = \log p(\mathbf{x} | \mathbf{z}_q, \theta) - \log K$$



Zhao Y. et al. Feature Quantization Improves GAN Training, 2020

Oord A., Vinyals O., Kavukcuoglu K. Neural Discrete Representation Learning, 2017

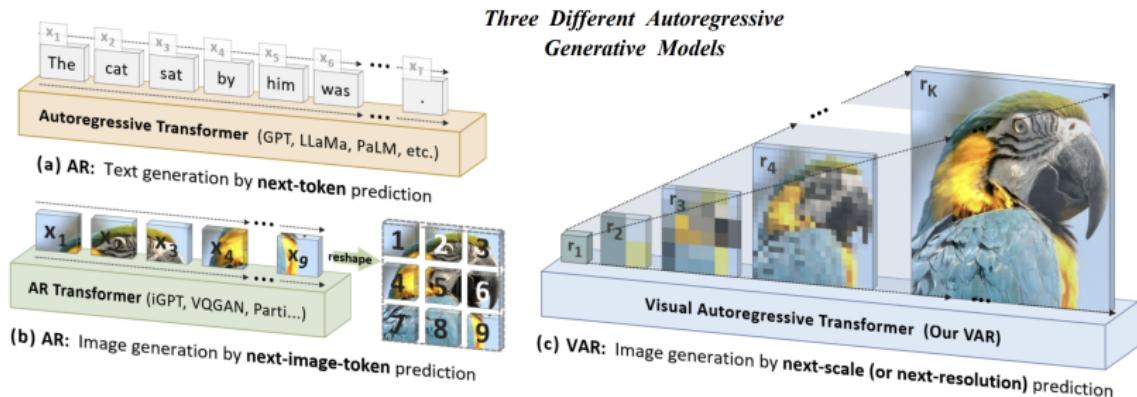
# Vector Quantized GAN



# LlamaGen

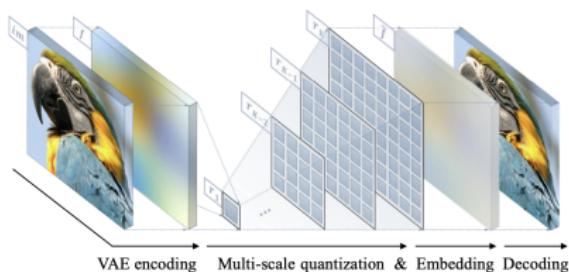


# Visual Autoregressive Modeling (VAR)

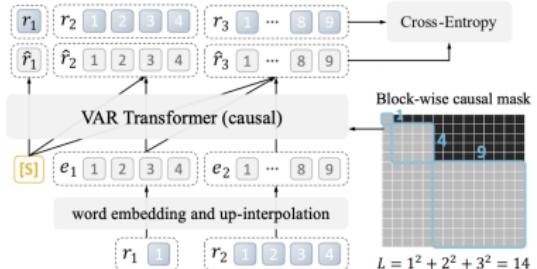


# Visual Autoregressive Modeling (VAR)

**Stage 1: Training multi-scale VQVAE on images**  
(to provide the ground truth for training Stage 2)



**Stage 2: Training VAR transformer on tokens**  
( $[S]$  means a start token with condition information)



# Outline

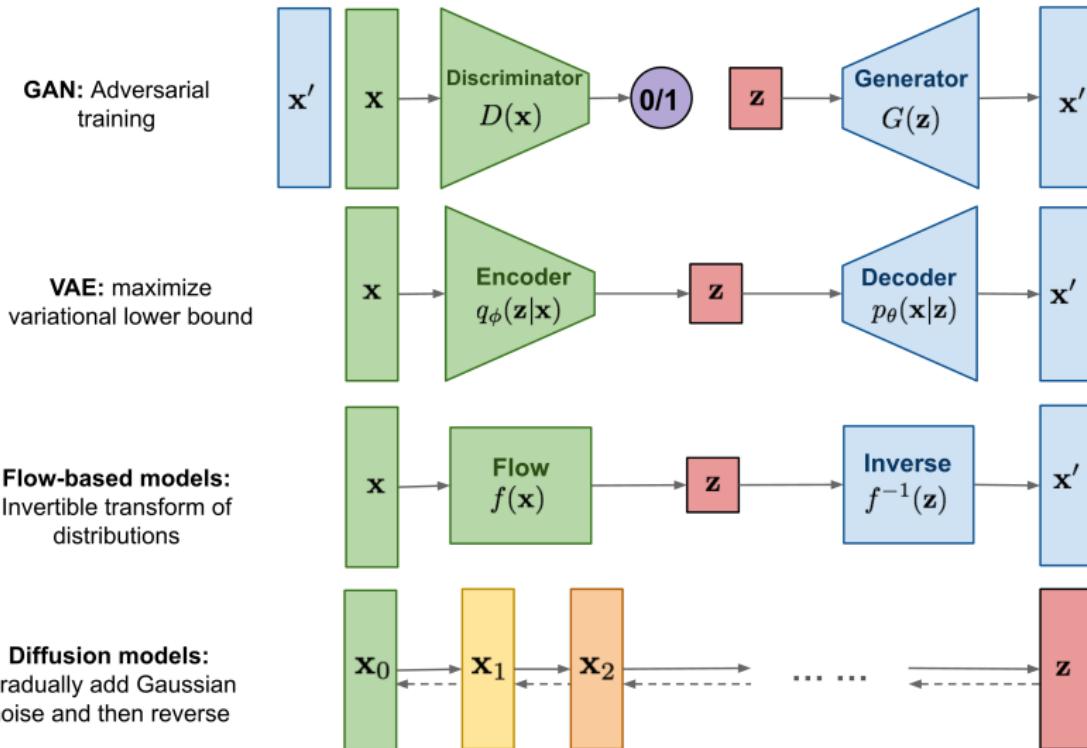
## 1. Latent models

Score-based models

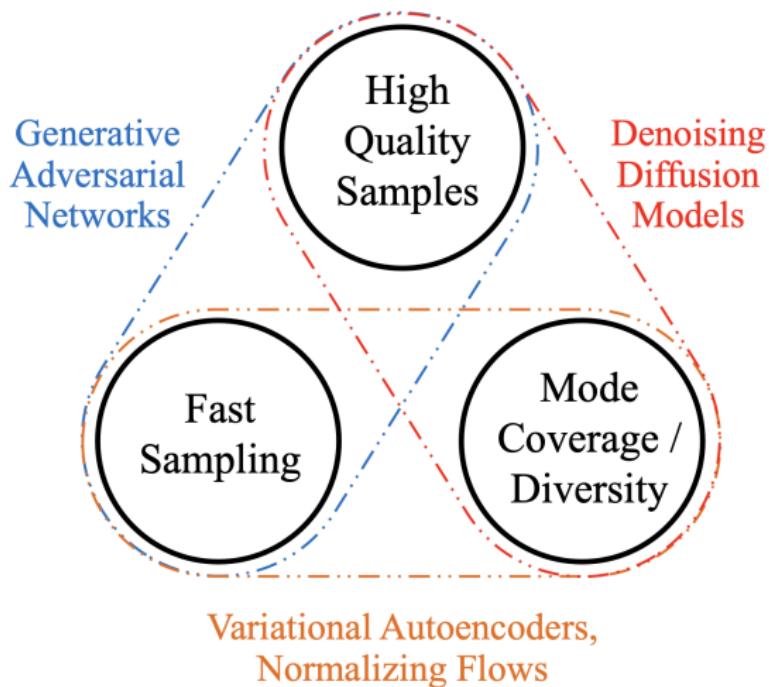
Autoregressive models

## 2. The worst course overview

# The worst course overview :)



# The worst course overview :)



# The worst course overview :)

Model	Efficient	Sample quality	Coverage	Well-behaved latent space	Disentangled latent space	Efficient likelihood
GANs	✓	✓	✗	✓	?	n/a
VAEs	✓	✗	?	✓	?	✗
Flows	✓	✗	?	✓	?	✓
Diffusion	✗	✓	?	✗	✗	✗

# Summary

