Generative Models
(This time without GANs)

June 3rd, 2021

Slides & Figures credits:
Assaf Shocher’s Slides, Introduction to Computer Vision 2020
Objective - Reminder

**Goal**: Find $\theta$ s.t. $\log \left( p_\theta (x^{(i)}) \right)$ is as high as possible
Latent Space Mapping Approach - Reminder
GANs - Reminder

\[
\mathcal{L}_{GAN} = \min_G \max_D \mathbb{E}_{x \sim p_{data}} \left[ \log D(x) \right] + \mathbb{E}_{z \sim p_z} \left[ \log(1 - D(G(z))) \right]
\]
Pros & Cons - GANs

• Pros:
  • Very high quality results
  • Fast inference

• Cons:
  • Difficult training
  • Evaluation is problematic
    • Using auxiliary methods: FID, IS
  • Mode collapse
Today – Generating Without GANs

• Autoencoders
• Variational Autoencoders
• VQ-VAE + VA-VAE2
• IMLE
Autoencoders

AE does not transform one *pre-determined* distribution to another!
Variational Autoencoders (Kingma&Welling 14’)

Encourage $p(z) \sim \mathcal{N}(0,1)$ by KL divergence:

$$\sum_{i=1}^{n} \sigma_i^2 + \mu_i^2 - \log(\sigma_i) - 1$$
Also check out the scale!
Probabilistic Interpretation

**Goal**: make \( \log(p_{\theta}(x^{(i)})) \) as high as possible

Using Bayes rule:

\[
p_{\theta}(x) = \frac{p_{\theta}(x|z)p_{\theta}(z)}{p_{\theta}(z|x)} \sim \mathcal{N}(0, 1)
\]

Sample \( z \) from \( z \mid x \sim \mathcal{N}(\mu_z|x, \Sigma_z|x) \)

Sample \( x \mid z \sim \mathcal{N}(\mu_x|z, \Sigma_x|z) \)
Probabilistic Interpretation

\[ \log p_\theta(x^{(i)}) = \text{Jacobian} \]

VAEs optimize a lower bound of the log likelihood

Slide credit: Stanford cs231n
Generate Data
Vector-Quantized (VQ) VAE (van den-Oord 2018)
Vector-Quantized (VQ) VAE

\[ \mathcal{L}_{\text{rec}} = \| \hat{x} - x \|_2^2 \quad \mathcal{L}_{\text{commit}} = \| E(x) - \text{sgn}(Q(x)) \|_2^2 \quad \mathcal{L}_{\text{codebook}} = \| \text{sgn}(E(x)) - Q(x) \|_2^2 \]

Quantization is non-differentiable!
VQ-VAE Reconstructions

Real  Reconstructed
Sampling New Instances

Training an autoregressive model to find $p(z)$

AutoRegressive Model (e.g. RNN)
Sampling New Instances

Once trained, use it to generate new decoder inputs

AutoRegressive Model (e.g. RNN)
Sampling New Instances - Results
VQ-VAE2
(Razavi & Van den-Oord 2019)

Codebook 1

Codebook 2

Global Structure

Fine details and textures
VQ-VAE2 – Sampling New Instances

Training two autoregressive models to find $p(z)$
VQ-VAE2 – Generation

Once both trained, use them to generate new decoder inputs
VQ-VAE2 – Results
VQ-VAE2 – Results
(\sim)VQ-VAE + Transformers: DALL-E

Ramesh et al., 2021

Teapot in the shape of a rubik’s cube

Soap-dispenser in the shape of a doughnut

Store front with ‘pytorch’
Implicit Maximum Likelihood Estimation (IMLE)
Li&Malik 2018

• GAN: All outputs should fool the discriminator
• Prone to mode collapse
Squares are data examples; circles are samples.

After training discriminator: highlighted regions are classified as real.

Generative Adversarial Nets

Generative Adversarial Nets

Generative Adversarial Nets

Generative Adversarial Nets

Generative Adversarial Nets

Generative Adversarial Nets

Generative Adversarial Nets

GANs only care about making each *sample* similar to some *data example*; it does not care about whether each *data example* is similar to some *sample*.


1809.09087, 2018
Implicit Maximum Likelihood Estimation (IMLE)
Li&Malik 2018

• GAN: All outputs should fool the discriminator
  • Close to some data point!
• IMLE: All data points should have close outputs

Avoiding mode collapse, easier training
IMLE - Method

Squares are data examples; circles are samples.

K. Li and J. Malik. Implicit Maximum Likelihood Estimation. arXiv:

1809.09087, 2018
Find the nearest sample to each data example.


1809.09087, 2018
Nearest samples are found.

IMLE - Method

Pull sample towards data example.

IMLE - Results
Summary

Alternatives to GANs (overcoming some of the problems):
• Autoencoders
• Variational Autoencoders
• VQ-VAE
• IMLE
Main principle:
In each scale, output patches belong to the distribution of input patches.
Diverse images generated from a single image

Source Image

Ours

SinGAN
QUESTIONS?