AICIP RESEARCH

ECE 599/692 – Deep Learning

Lecture 11 – Generative Adversarial Networks

Hairong Qi, Gonzalez Family Professor
Electrical Engineering and Computer Science
University of Tennessee, Knoxville
http://www.eecs.utk.edu/faculty/qi
Email: hqi@utk.edu

AICIP RESEARCH

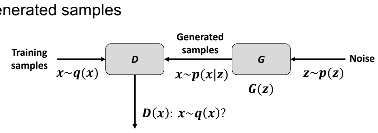
Outline

- Lecture 11: GAN – Introduction and theoretic analysis
- Lecture 12: Theoretic analysis and other generative models – VAE
- Lecture 13: Conditional GAN
- Lecture 14: Implementation

AICIP RESEARCH

GAN

- Two neural networks compete against each other
  - A generator network G: mimic training samples to fool the discriminator
  - A discriminator network D: discriminate training samples and generated samples



For *D*: $\max_D \ \mathbb{E}_{x\sim q(x)}[\log(D(x))] + \mathbb{E}_{z\sim p(z)}[\log(1 - D(G(z)))]$

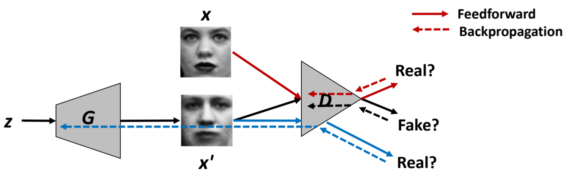
For *G*: $\min_G \ \mathbb{E}_{z\sim p(z)}[\log(1 - D(G(z)))]$

GAN

AICIP RESEARCH

The objective function of GANs:

$$\min_G \max_D \ \mathbb{E}_{x\sim q(x)}[\log(D(x))] + \mathbb{E}_{z\sim p(z)}[\log(1 - D(G(z)))]$$



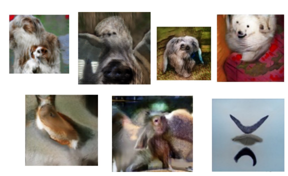
THE UNIVERSITY OF TENNESSEE

4

GAN - Drawbacks

AICIP RESEARCH

- Mode missing problem
- Generate unrealistic images
- Hard to learn to generate discrete data, e.g., text

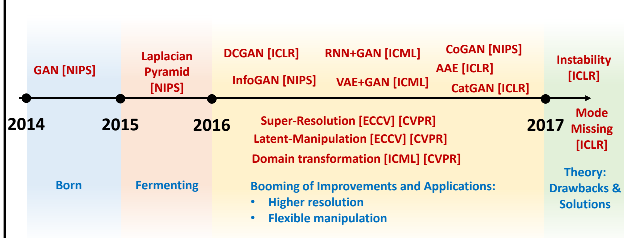


THE UNIVERSITY OF TENNESSEE

5

Evolution of GAN

AICIP RESEARCH



THE UNIVERSITY OF TENNESSEE

6

AICIP RESEARCH

Image In-painting [13,14]

(a) Input context (b) Human artist (c) Context Encoder ($L2$ loss) (d) Context Encoder ($L2$ + Adversarial loss)

[13] Pathak, Deepak, et al. "Context encoders: Feature learning by inpainting." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016.
[14] Li, Yijun, et al. "Generative Face Completion." *CVPR* (2017).

THE UNIVERSITY OF TENNESSEE KNOXVILLE

AICIP RESEARCH

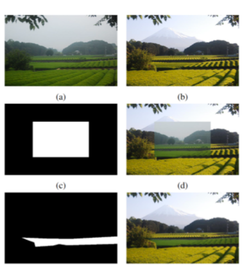
Image Editing

[11] Larsen, Anders Boesen Lindbo, et al. "Autoencoding beyond pixels using a learned similarity metric."ICML(2016).

THE UNIVERSITY OF TENNESSEE KNOXVILLE

AICIP RESEARCH

Image Blending [16]



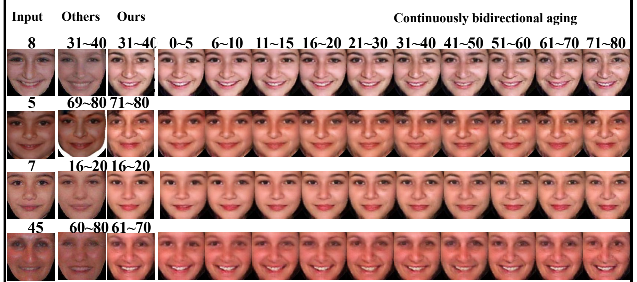
(a) (b) (c) (d) (e) (f)

[16]Wu, Huikai, et al. "GP-GAN: Towards Realistic High-Resolution Image Blending." *arXiv preprint arXiv:1703.07195* (2017).

THE UNIVERSITY OF TENNESSEE KNOXVILLE

Age Progression and Regression

AICIP RESEARCH



Project page: https://zzutk.github.io/Face-Aging-CAAE

THE UNIVERSITY OF TENNESSEE KNOXVILLE

10

THE UNIVERSITY OF TENNESSEE KNOXVILLE

AICIP RESEARCH

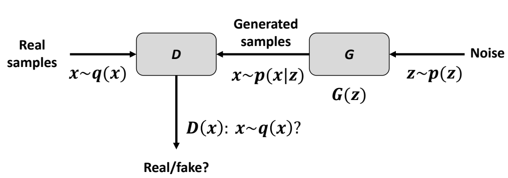
GAN – Theoretic Reasoning

11

Theoretic Reasoning

AICIP RESEARCH

- Instability of GAN
  - Mode missing
  - Gradient vanishing



THE UNIVERSITY OF TENNESSEE KNOXVILLE

Instability – Mode Missing

$$\min_G \max_D \ \mathbb{E}_{x\sim q(x)}[\log(D(x))] + \mathbb{E}_{z\sim p(z)}[\log(1 - D(G(z)))]$$

Expectation: $\mathbb{E}_{x\sim q(x)}[p(x)] = \int_x q(x)p(x)\,dx$

$$\min_G \max_D \int_x q(x)\log(D(x))\,dx + \int_z p(z)\log(1 - D(G(z)))\,dz$$

Usually, G and D are updated alternatively

Instability – Mode missing (cont'd)

Fix *G*, $\min_G \max_D \int_x q(x)\log(D(x))\,dx + \int_z p(z)\log(1 - D(G(z)))\,dz$

$$= \max_D \int_x q(x)\log(D(x)) + p_g(x)\log(1 - D(x))\,dx$$

$$D^*(x) = \frac{q(x)}{q(x) + p_g(x)}$$

Fix D^*,

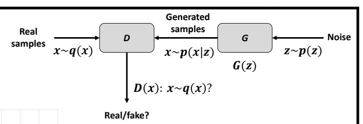
$$= \min_G \int_x p_g(x)\log\left(1 - \frac{q(x)}{q(x) + p_g(x)}\right)dx$$

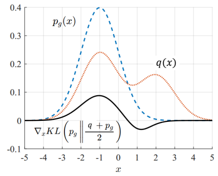
$$= \min_G \int_x p_g(x)\log\left(\frac{p_g(x)}{q(x) + p_g(x)}\right)dx$$

$$= \min_G \ D_{KL}(p_g || q + p_g)$$

$$= \min_G \ D_{KL}(p_g || \frac{q + p_g}{2}) - 2\log 2$$

Instability – Mode missing (cont'd)





$q(x) < p_g(x)$: Unrealistic samples
$q(x) > p_g(x)$: Mode missing

- **Punish more on generating unrealistic samples**
- **Punish less on mode missing**

Instability – Gradient vanishing

AICIP RESEARCH

$$\min_G \max_D \ \mathbb{E}_{x\sim q(x)}[\log(D(x))] + \mathbb{E}_{z\sim p(z)}[\log(1 - D(G(z)))]$$

$$\begin{aligned}\nabla_\theta \mathcal{L}_G =& \nabla_\theta \mathbb{E}_{z\sim p_z}[\log(1 - D(G_\theta(z)))] \\ =& \mathbb{E}_{z\sim p_z}\left[\frac{1}{D(G_\theta(z)) - 1}\frac{\partial D(G_\theta(z))}{\partial G_\theta(z)}\frac{\partial G_\theta(z)}{\partial \theta}\right] \\ =& \mathbb{E}_{x\sim p_g}\left[\frac{\nabla_x D(x)\nabla_\theta G_\theta}{D(x) - 1}\right],\end{aligned}$$

where $\|\mathbb{E}_{x\sim p_g}[\nabla_\theta G_\theta]\|_2$ is bounded for a differentiable generator. If D is a perfect discriminator $D^\dagger$, $D^\dagger(x)|_{x\sim p_x} = 1$ and $D^\dagger(x)|_{x\sim p_g} = 0$. Note that a perfect discriminator is an optimal discriminator, but an optimal discriminator is not necessarily a perfect one.

AICIP RESEARCH

When D is approaching $D^\dagger$ and $x \sim p_g$,

$$\lim_{D\to D^\dagger} D(x) = 0,$$

$$\lim_{D\to D^\dagger} \nabla_x D(x) = 0.$$

Therefore,

$$\lim_{D\to D^\dagger} \mathbb{E}_{x\sim p_g}\left[\frac{\nabla_x D(x)\nabla_\theta G_\theta}{D(x) - 1}\right] = 0,$$

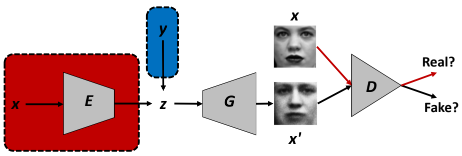
Instability

AICIP RESEARCH

- If “D” is optimal, it will cause gradient vanishing for G.
- If “D” is poor, the gradient of G is unstable, huge occlusion.

Don’t train “D” too good or too poor!

GAN-based Image Manipulation



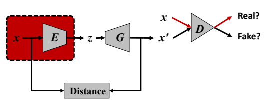
- Seldom use original GAN
- Concatenate an encoder to *G*
- Concatenate extra feature to *z*

Stabilizing GAN – Adding an AE

In GAN, the generated distribution, $p_g(x)$, is matched to the distribution specified by D, rather than the real distribution

In GAN, $D(x|p_g) = D(x|q)$, which is indirect matching.

Ideally, $p_g(x) = q(x)$, which is direct matching.



without AE: $\mathbb{E}_{x\sim p_x}[\log D(x)] + \mathbb{E}_{z\sim p_z}[\log(1 - D(G(z)))]$

with AE: $\mathbb{E}_{x\sim p_x}[\log (D(x)(1 - D(\mathcal{H}(x)))) + \lambda\mathcal{L}(x, \mathcal{H}(x))]$

where λ balances the effect of reconstruction error, and $\mathcal{H}(x) = G(E(x))$.

Stabilizing GAN - adding an AE (cont'd)

$$\mathbb{E}_{x\sim p_x}[\log (D(x)(1 - D(\mathcal{H}(x)))) + \lambda\mathcal{L}(x, \mathcal{H}(x))]$$

Remedy to Mode Missing

$$\begin{aligned} KL(p_x\|p_g) &= \int_{\Omega_x} p_x(x)\log\frac{p_x(x)}{p_g(x)}dx \\ &= \int_{\Omega_x} p_x(x)\log p_x(x)dx - \int_{\Omega_x} p_x(x)\log p_g(x)dx \\ &= \underbrace{\mathbb{E}_{x\sim p_x}[\log p_x(x)]}_{\text{Constant because } p_x(x) \text{ is fixed given a training set.}} - \underbrace{\mathbb{E}_{x\sim p_x}[\log p_g(x)]}_{\text{Equivalent to fitting generated samples to real ones}}, \end{aligned}$$

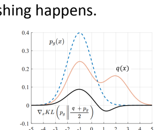
Remedy to Unrealistic Generation

Remedy to Gradient Vanishing

The gradient of AE part

$$\lambda\nabla_\theta\mathcal{L}(x, \mathcal{H}(x))$$

is larger than 0 when gradient vanishing happens.



$q(x) < p_g(x)$: Unrealistic samples
$q(x) > p_g(x)$: Mode missing

Acknowledgement

AICIP RESEARCH

Most slides are taken from student presentations at AICIP group meetings, including mainly those from Zhifei Zhang.

THE UNIVERSITY OF TENNESSEE KNOXVILLE

22