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

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

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

The objective function of GANs:

\[
\min_{\mathcal{G}} \max_{\mathcal{D}} \mathbb{E}_{x \sim q(x)} [\log \mathcal{D}(x)] + \mathbb{E}_{z \sim \mathcal{P}(z)} [\log (1 - \mathcal{D}(\mathcal{G}(z)))]
\]

GAN - Drawbacks

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

Evolution of GAN

- Born: 2014
- Fermenting: 2015
- Booming: 2016
- Instability: 2017

- Theory: [ICLR]
- Drawbacks & Solutions: [ICLR]
- Mode missing: [ICLR]
- Instability: [ICLR]
- Latent manipulation: [ICLR] [CVPR]
- Domain transformation: [ICML] [CVPR]
- Higher resolution
- Flexible manipulation
Image In-painting [11,14]

Image Editing

Image Blending [16]


Age Progression and Regression

<table>
<thead>
<tr>
<th>Input</th>
<th>Others</th>
<th>Ours</th>
<th>Continuously bidirectional aging</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>31-40</td>
<td>31-40</td>
<td>0-5</td>
</tr>
<tr>
<td>5</td>
<td>60-80</td>
<td>71-80</td>
<td>6-10</td>
</tr>
<tr>
<td>7</td>
<td>16-29</td>
<td>16-29</td>
<td>11-15</td>
</tr>
<tr>
<td>45</td>
<td>60-80</td>
<td>61-70</td>
<td>21-30</td>
</tr>
</tbody>
</table>

Project page: [https://zzuk.github.io/Face-Aging-CAAE](https://zzuk.github.io/Face-Aging-CAAE)

GAN – Theoretic Reasoning

Theoretic Reasoning

- Instability of GAN
  - Mode missing
  - Gradient vanishing

```
D: x ~ q(x)
G: z ~ p(z)
D(z): x ~ q(x)?
```

For D:

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

For G:

\[
\min \mathbb{E} z \sim p(z) \log (1 - D(G(z)))
\]
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: \( \mathbb{E}_{x \sim q(x)}[p(x)] = \int_x q(x)p(x) \, dx \)

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

Usually, \( G \) and \( D \) are updated alternatively.

---

Instability – Mode missing (cont’d)

Fix \( G \),

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

\[ = \max_D \int_x q(x) \log(D(x)) + p_G(z) \log(1 - D(x)) \, dx \]

\[ D^*(x) = \frac{q(x)}{q(x) + p_G(x)} \]

Fix \( D^* \),

\[ = \min_G \int_x p_G(z) \log \left(1 - \frac{q(x)}{q(x) + p_G(x)} \right) \, dz \]

\[ = \min_G \int_x p_G(z) \log \left(\frac{p_G(z)}{q(x) + p_G(x)} \right) \, dz \]

\[ = \min_G \mathcal{D}_G(p_G || 1 + p_G) \]

\[ = \min_G \mathcal{D}_G(p_G || 1 + p_G) - 2 \log 2 \]

---

Instability – Mode missing (cont’d)

• Punish more on generating unrealistic samples
• Punish less on mode missing
Instability – Gradient vanishing

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

\[ \nabla_L G = \nabla \mathbb{E}_{x \sim p(x)} \left[ \log(1 - D(G(z))) \right] \]

\[ = \mathbb{E}_{x \sim p(x)} \left[ \nabla _x D(x) \frac{1}{D(x) - 1} \frac{\partial D(G(x))}{\partial G(x)} \frac{\partial G(x)}{\partial \theta} \right] \]

\[ = \mathbb{E}_{x \sim p(x)} \left[ \nabla _x D(x) \frac{1}{D(x) - 1} \frac{\partial D(G(x))}{\partial G(x)} \right] \]

where \( \| \mathbb{E}_{x \sim p(x)} \nabla G(x) \|_2 \) is bounded for a differentiable generator. If \( D \) is a perfect discriminator \( D^1, D^1(x)\vert_{x \sim p(x)} = 1 \) and \( D^1(x)\vert_{x \sim p(x)} = 0 \). Note that a perfect discriminator is an optimal discriminator, but an optimal discriminator is not necessarily a perfect one.

Instability

- 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!
Stabilizing GAN by Incorporating An Autoencoder

In GAN, the generated distribution, \( p_g(x) \), is matched to the distribution specified by \( D \), rather than the real distribution. Ideally, \( p_g(x) = q(x) \), which is direct matching.

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

With AE:
\[
\mathbb{E}_{x \sim p_D}[\log (D(x)(1 - D(H(z)))) + \lambda \mathbb{L}(x; H(z))]
\]

where \( \lambda \) balances the effect of reconstruction error, and \( H(z) = G(E(z)) \).

Stabilizing GAN - Adding an AE (cont’d)

\[
\mathbb{E}_{x \sim p_D}[\log (D(x)(1 - D(H(z)))) + \lambda \mathbb{L}(x; H(z))]
\]

Remedy to Mode Missing

\[
\mathcal{L}(x; H(z)) = \int p(z|x) \log \frac{p(z|x)}{H(z)} dz
\]

Remedy to Gradient Vanishing

The gradient of AE part
\[
\nabla_x \mathbb{L}(x; H(z))
\]
is larger than 0 when gradient vanishing happens.

Constant because \( p(z|x) \) is fixed given a training set.

Equivalent to fitting generated samples to real ones.

Remedy to Unrealistic Generation

\( q(z|x) \) unrealistic samples
\( q(z|x) \) Mode missing
Acknowledgement

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