The power of two samples in generative adversarial networks

Sewoong Oh

Department of Industrial and Enterprise Systems Engineering
University of Illinois at Urbana-Champaign
Generative models learn fundamental representations.

\[ G^{-1}(\hat{g}) + Z_{\text{glasses}} = Z_{\text{glasses}} \]

[DCCGAN, Radford et al. 2015]
Generative Adversarial Networks (GAN)

\[ G(Z) \rightarrow X \]

\[ \text{Real data} \]

\[ \text{Generator } G(Z) \]

\[ \text{Discriminator } D(X) \]

\[ \min_G \max_D V(G, D) \]
Challenges in training GAN

1. Instability: non-convergence of training loss

2. Evaluation: likelihood is not available

3. Mode collapse: loss of diversity
“Mode collapse” is a main challenge
“Mode collapse” is a main challenge
“Mode collapse” is a main challenge

- “A man in an orange jacket with sunglasses and a hat ski down a hill.”

- “This guy is in black trunks and swimming underwater.”

- “A tennis player in a blue polo shirt is looking down at the green court.”

[“Generating interpretable images with controllable structure”, by Reed et al., 2016]
Lack of diversity is easier to detect if the discriminator sees multiple samples jointly.

- "Distributional Adversarial Networks", Li, Alvarez-Melis, Xu, Jegelka, Sra, 2017
New framework: PacGAN

- lightweight overhead
- experimental results
- principled
Benchmark tests

<table>
<thead>
<tr>
<th>Modes</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAN</td>
<td>17.3</td>
</tr>
<tr>
<td>PacGAN2</td>
<td>23.8</td>
</tr>
<tr>
<td>PacGAN3</td>
<td>24.6</td>
</tr>
<tr>
<td>PacGAN4</td>
<td>24.8</td>
</tr>
</tbody>
</table>
Benchmark datasets from VEEGAN paper

<table>
<thead>
<tr>
<th>Real data</th>
<th>DCGAN</th>
<th>PacDCGAN2</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Real data" /></td>
<td><img src="image2" alt="DCGAN" /></td>
<td><img src="image3" alt="PacDCGAN2" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Modes (Max 1000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCGAN</td>
</tr>
<tr>
<td>ALI</td>
</tr>
<tr>
<td>Unrolled GAN</td>
</tr>
<tr>
<td>VEEGAN</td>
</tr>
<tr>
<td>PacDCGAN2</td>
</tr>
<tr>
<td>PacDCGAN3</td>
</tr>
<tr>
<td>PacDCGAN4</td>
</tr>
</tbody>
</table>
Intuition behind packing via toy example

Target distribution $P$

Generator $Q_1$
with mode collapse

Generator $Q_2$
without mode collapse

$d_{TV}(P, Q_1) = 0.2$
d$d_{TV}(P, Q_2) = 0.2$
Intuition behind packing via toy example

Target distribution $P$

Generator $Q_1$
with mode collapse

Generator $Q_2$
without mode collapse

Target distribution $P$

$P \times P$

$Q_1 \times Q_1$

$Q_2 \times Q_2$

$P \times P, Q_1 \times Q_1 = 0.36$

$P \times P, Q_2 \times Q_2 = 0.24$

$\text{d}_{TV}(P^2, Q_1^2)$

$\text{d}_{TV}(P^2, Q_2^2)$
Intuition behind packing via toy example

Target distribution $P$

Generator $Q_1$ with mode collapse

Generator $Q_2$ without mode collapse

$d_{TV}(P \times P, Q_1 \times Q_1) = 0.36$

$d_{TV}(P \times P, Q_2 \times Q_2) = 0.24$
Evolution of TV distances

Total variation $d_{TV}(P^m, Q^m)$

Through packing, the target-generator pairs are expanded over the strengths of the mode collapse.
Evolution of TV distances through the prism of packing

Through packing, the target-generator pairs are expanded over the strengths of the mode collapse.
\[ d_{TV}(P^m, Q^m) \]

\[
\begin{align*}
\max_{P, Q} / \min_{P, Q} & \quad d_{TV}(P^2, Q^2) \\
\text{subject to} & \quad d_{TV}(P, Q) = \tau
\end{align*}
\]

- we focus on \( m = 2 \) for this talk
- this is easy, but we have a new proof technique
- nothing to do with mode collapse, but we use it as proof technique
We say a pair \((P, Q)\) of a target distribution \(P\) and a generator distribution \(Q\) has \((\varepsilon, \delta)\)-mode collapse if there exists a set \(S\) such that

\[
P(S) \geq \delta, \quad \text{and} \quad Q(S) \leq \varepsilon.
\]
We say a pair \((P, Q)\) of a target distribution \(P\) and a generator distribution \(Q\) has \((\varepsilon, \delta)\)-mode collapse if there exists a set \(S\) such that

\[ P(S) \geq \delta, \quad \text{and} \quad Q(S) \leq \varepsilon. \]
We say a pair \((P, Q)\) of a target distribution \(P\) and a generator distribution \(Q\) has \((\varepsilon, \delta)\)-mode collapse if there exists a set \(S\) such that

\[
P(S) \geq \delta, \quad \text{and} \quad Q(S) \leq \varepsilon.
\]
We say a pair \((P, Q)\) of a target distribution \(P\) and a generator distribution \(Q\) has \((\varepsilon, \delta)\)-mode collapse if there exists a set \(S\) such that

\[
P(S) \geq \delta, \quad \text{and} \quad Q(S) \leq \varepsilon.
\]

**Definition [mode collapse region]**

\[
\mathcal{R}(P, Q) = \{ (\varepsilon, \delta) : P(S) \geq \delta, Q(S) \leq \varepsilon \}
\]

**Intuition from Blackwell**

**Target distribution \(P\)**

\[
P = \begin{cases} 
1 & \text{at 1} \\
\delta & \text{at \(\delta\)} \\
1 & \text{at 1}
\end{cases}
\]

**Generator \(Q_1\)**

\[
Q_1 = \begin{cases} 
1 & \text{at 1} \\
1.25 & \text{at 1.25} \\
\varepsilon & \text{at \(\varepsilon\)} \\
0.2 & \text{at 0.2} \\
1 & \text{at 1}
\end{cases}
\]

**\(\mathcal{R}(P, Q_1)\)**

\[
\mathcal{R}(P, Q_1) = \{ (\varepsilon, \delta) : P(S) \geq \delta, Q(S) \leq \varepsilon \}
\]
Intuition from Blackwell

**Definition [mode collapse region]**

We say a pair \((P, Q)\) of a target distribution \(P\) and a generator distribution \(Q\) has \((\varepsilon, \delta)\)-**mode collapse** if there exists a set \(S\) such that

\[
P(S) \geq \delta \quad \text{and} \quad Q(S) \leq \varepsilon.
\]
Intuition from Blackwell

Definition [mode collapse region]

We say a pair \((P, Q)\) of a target distribution \(P\) and a generator distribution \(Q\) has \((\varepsilon, \delta)\)-mode collapse if there exists a set \(S\) such that

\[
P(S) \geq \delta \quad \text{and} \quad Q(S) \leq \varepsilon.
\]
Intuition from Blackwell

**Definition [mode collapse region]**

We say a pair \((P, Q)\) of a target distribution \(P\) and a generator distribution \(Q\) has \((\varepsilon, \delta)\)-**mode collapse** if there exists a set \(S\) such that

\[
P(S) \geq \delta \quad \text{and} \quad Q(S) \leq \varepsilon.
\]

![Diagram showing target distribution \(P\) and generator distribution \(Q_1\) with mode collapse](image)
We say a pair \((P, Q)\) of a target distribution \(P\) and a generator distribution \(Q\) has \((\varepsilon, \delta)\)-mode collapse if there exists a set \(S\) such that

\[
P(S) \geq \delta, \quad \text{and} \quad Q(S) \leq \varepsilon.
\]
Definition [mode collapse region]

We say a pair \((P, Q)\) of a target distribution \(P\) and a generator distribution \(Q\) has \((\varepsilon, \delta)\)-mode collapse if there exists a set \(S\) such that

\[
P(S) \geq \delta, \quad \text{and} \quad Q(S) \leq \varepsilon.
\]
Intuition from Blackwell

**Definition [mode collapse region]**

We say a pair \((P, Q)\) of a target distribution \(P\) and a generator distribution \(Q\) has \((\varepsilon, \delta)\)-**mode collapse** if there exists a set \(S\) such that

\[
P(S) \geq \delta \quad \text{and} \quad Q(S) \leq \varepsilon.
\]

The image shows a target distribution \(P\) and a generator distribution \(Q_2\) without mode collapse. The distance between \(P\) and \(Q_2\) is denoted as \(d_{TV}(P, Q_2) = 0.2\).
Upper bound

\[
\begin{align*}
\max_{P,Q} & \quad d_{TV}(P^2, Q^2) \\
\text{subject to} & \quad d_{TV}(P, Q) = \tau \\
\end{align*}
\]

Blackwell's theorem

\[
R(P, Q) \subseteq R(P', Q') \implies R(P^2, Q^2) \subseteq R(P'^2, Q'^2)
\]
Upper bound

\[
\max_{P,Q} d_{TV}(P^2, Q^2)
\]

subject to

\[
d_{TV}(P, Q) = \tau
\]

\[
\mathcal{R}(P, Q) \subseteq \mathcal{R}_{outer}(\tau)
\]

Blackwell's theorem

\[
\mathcal{R}(P, Q) \subseteq \mathcal{R}(P', Q') \Rightarrow \mathcal{R}(P^2, Q^2) \subseteq \mathcal{R}(P'^2, Q'^2)
\]
Upper bound

\[
\max_{P,Q} \quad d_{TV}(P^2, Q^2)
\]

subject to

\[d_{TV}(P, Q) = \tau\]

\[\mathcal{R}(P, Q) \subseteq \mathcal{R}_{outer}(\tau)\]
**Upper bound**

\[
\begin{align*}
\max_{P,Q} & \quad d_{TV}(P^2, Q^2) \\
\text{subject to} & \quad d_{TV}(P, Q) = \tau \\
\mathcal{R}(P, Q) & \subseteq \mathcal{R}_{\text{outer}}(\tau) \\
\mathcal{R}(P^2, Q^2) & \subseteq \mathcal{R}(P^2_{\text{outer}}, Q^2_{\text{outer}})
\end{align*}
\]

**Blackwell’s theorem**

\[
\begin{align*}
\mathcal{R}(P, Q) & \subseteq \mathcal{R}(P', Q') \\
\Rightarrow & \quad \mathcal{R}(P^2, Q^2) \subseteq \mathcal{R}(P'^2, Q'^2)
\end{align*}
\]
Upper bound

\[
\max_{P, Q} d_{TV}(P^2, Q^2)
\]
subject to
\[
d_{TV}(P, Q) = \tau
\]

\[
\mathcal{R}(P, Q) \subseteq \mathcal{R}_{outer}(\tau)
\]
\[
\mathcal{R}(P^2, Q^2) \subseteq \mathcal{R}(P_{outer}^2, Q_{outer}^2)
\]
\[
d_{TV}(P^2, Q^2) \leq d_{TV}(P_{outer}^2, Q_{outer}^2) \left(1 - (1 - \tau)^2\right)
\]

Blackwell’s theorem

\[
\mathcal{R}(P, Q) \subseteq \mathcal{R}(P', Q')
\]
\[
\Rightarrow \mathcal{R}(P^2, Q^2) \subseteq \mathcal{R}(P'^2, Q'^2)
\]
$d_{TV}(P^m, Q^m)$

**max / min** $d_{TV}(P^2, Q^2)$

subject to $d_{TV}(P, Q) = \tau$
PacGAN naturally penalizes mode collapse

\[ d_{TV}(P^m, Q^m) \]

without \((\varepsilon_0, \delta_0)\)-mode collapse

\[
\max_{P,Q} \quad d_{TV}(P^2, Q^2) \\
\text{s.t.} \quad d_{TV}(P, Q) = \tau \\
\text{no} \ (\varepsilon_0, \delta_0)\)-mode collapse
\]
PacGAN naturally penalizes mode collapse

\[ d_{TV}(P^m, Q^m) \]

without \((\varepsilon_0, \delta_0)\)-mode collapse

\[
\max_{P, Q} d_{TV}(P^2, Q^2) \\
\text{s.t.} \quad d_{TV}(P, Q) = \tau \\
\text{no} \ (\varepsilon_0, \delta_0)\text{-mode collapse}
\]
PacGAN naturally penalizes mode collapse

without \((\varepsilon_0, \delta_0)\)-mode collapse

\[
\max_{P,Q} \quad d_{\text{TV}}(P^m, Q^m)
\]
\[
\text{s.t.} \quad d_{\text{TV}}(P, Q) = \tau
\]
no \((\varepsilon_0, \delta_0)\)-mode collapse

with \((\varepsilon_1, \delta_1)\)-mode collapse

\[
\min_{P,Q} \quad d_{\text{TV}}(P^m, Q^m)
\]
\[
\text{s.t.} \quad d_{\text{TV}}(P, Q) = \tau
\]
\((\varepsilon_1, \delta_1)\)-mode collapse

\[
\frac{\varepsilon_1}{\varepsilon_0} = \frac{\delta_1}{\delta_0}
\]
Size of the discriminator

# modes captured

# of parameters in $D(\cdot)$

Mini-batch discrimination requires +38,748,557, PacGAN2 requires +54
0-1 loss (Total Variation) vs. cross entropy loss (Jensen-Shannon Divergence)

Jensen-Shannon is better measure for detecting mode collapse
Our paper is available at: https://arxiv.org/abs/1712.04086

All codes for the experiments at: https://github.com/fjxmlzn/PacGAN

Zinan Lin (CMU)  Ashish Khetan (UIUC)  Giulia Fanti (CMU)