Weakly supervised 3D Reconstruction with Adversarial Constraint

JunYoung Gwak, Christopher Choy, Manmohan Chandraker, Animesh Garg, Silvio Savarese
3D reconstruction

2D mask weak supervision + Adversarial constraint = Realistic 3D reconstruction of the image
3D reconstruction

**Full 3D supervision**
[Choy et al, Wu et al, Girdhar et al]

3D models are **expensive annotation**

**2D mask supervision**
[Yan et al.]

Capped by limitations of **visual hull**
Weakly supervised 3D Reconstruction

2D mask weak supervision

+ 

Adversarial constraint

= 

Realistic 3D reconstruction of the image

\[
\begin{align*}
\text{minimize} & \quad x \\
\text{subject to} & \quad \text{Reproduction Error}(x) \\
\text{Reconstruction} & \quad x \text{ to be a valid chair}
\end{align*}
\]
Adversarial Constraint

• Rewrite constrained optimization using **log barrier method**

• Train constraint function $g(x)$ as **GAN discriminator**

\[
\begin{align*}
\text{minimize} & \quad \text{ReprojectionError}(x) \\
\text{subject to} & \quad \text{Reconstruction } x \text{ to be a valid chair}
\end{align*}
\]

\[
\begin{align*}
\text{minimize} & \quad \text{ReprojectionError}(x) - \frac{1}{t} \log g(x) \\
\text{minimize} & \quad \mathbb{E}_{x^* \sim p} \log g(x^*) + \mathbb{E}_{\hat{x} \sim q} \log(1 - g(\hat{x}))
\end{align*}
\]
2D mask weak supervision

- Raytrace Pooling:
  Renders reconstruction to mask

\[
\begin{align*}
\text{minimize}_{x} & \quad \text{ReprojectionError}(x) \\
\text{subject to} & \quad \text{Reconstruction } x \text{ to be a valid chair}
\end{align*}
\]

\[
\text{minimize}_{x} \quad \text{ReprojectionError}(x) - \frac{1}{t} \log g(x)
\]
Result 1: Synthetic image reconstruction

Qualitative results

<table>
<thead>
<tr>
<th>Input</th>
<th>Ground-truth</th>
<th>2D supervision</th>
<th>2D supervision + Adversarial Constraint (ours)</th>
</tr>
</thead>
</table>
Result 1: Synthetic image reconstruction

Quantitative results

![Bar chart showing IOU for different numbers of views.](chart)

- **Voxel Carving**
- **2D supervision**
- **2D supervision + Adversarial Constraint (ours)**
Result 2: Real image reconstruction

**Single-view real image** reconstruction on ObjectNet3D

<table>
<thead>
<tr>
<th>Input</th>
<th>Ground-truth</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Input Image 1]</td>
<td>![Ground-truth Image 1]</td>
<td>![Ours Image 1]</td>
</tr>
<tr>
<td>![Input Image 2]</td>
<td>![Ground-truth Image 2]</td>
<td>![Ours Image 2]</td>
</tr>
<tr>
<td>![Input Image 3]</td>
<td>![Ground-truth Image 3]</td>
<td>![Ours Image 3]</td>
</tr>
</tbody>
</table>

**Multi-view real-image** reconstruction on Stanford Online Product

<table>
<thead>
<tr>
<th>Input</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Input Image 4]</td>
<td>![Ours Image 4]</td>
</tr>
<tr>
<td>![Input Image 5]</td>
<td>![Ours Image 5]</td>
</tr>
<tr>
<td>![Input Image 6]</td>
<td>![Ours Image 6]</td>
</tr>
</tbody>
</table>
Result 3: Hidden representation analysis

**Linear interpolation** of hidden variables of two images

**Semantic feature arithmetic** of hidden variables

hole in the back, four legs
solid back, four legs
hole in the back, two legs
solid back, two legs
Conclusion

- 2D mask weak supervision
  - Affordable and practical
  - Made possible by Raytrace Pooling
- Adversarial constraint
  - GAN as log barrier constraint
  - Overcomes limitation of visual hull

2D mask weak supervision + Adversarial constraint = Realistic 3D reconstruction of the image