**Image-to-Image Translation with Conditional GAN**

**Weini Yu, Zhouchangwan Yu, Jason Hu**

**STANFORD UNIVERSITY**

### Introduction
- Image-to-image translation are tasks that take in input images and generate or manipulate them into a different visual space. Traditionally this task requires hand-crafted machinery.
- In this study we explore image translation using **conditional-general adversarial networks (C-GAN)**, in which we translate images using GAN conditioned on input images and generate desired output image.

### Data
- Paired aerial images and corresponding maps scraped from Google Maps.
- Image size 600x600 in JPEG
- 1097 training, 1098 validation, and 1098 test examples.

### Preprocessing
- Resize to 256 x 256
- Random crop to 256 x 256
- Random horizontal flip
- Zero center and normalize all pixel values to [-1, 1]

### Method
**Conditional GAN**

**Conditional adversarial loss:**
\[
L_{cond}(G, D) = P_Y[D_y(G(x))] + P_X[D_x(1 - D_y(G(x)))]
\]

**Objective of minimax game:**
\[
(G^*, D^*) = \arg\min_G \arg\max_D \left( L_{cond}(G, D) + \lambda M_{c}(D) \right)
\]

**Figure 2:** Conditional-GAN model

### Results and Analysis
- Fig. 2: Generated map images of different architecture and hyperparameters

<table>
<thead>
<tr>
<th>Model</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>1.5661e²</td>
</tr>
<tr>
<td>U-Net</td>
<td>1.5058e²</td>
</tr>
<tr>
<td>ResNet-50</td>
<td>1.3298e²</td>
</tr>
<tr>
<td>ResNet-9</td>
<td>1.1909e²</td>
</tr>
<tr>
<td>ResNet50 (He et al.)</td>
<td>1.5845e²</td>
</tr>
</tbody>
</table>

**Table 1:** Mean Squared Error between different models and ground truth

**Figure 3:** Train set discriminator score

- **Fig. 4:** Loss of ResNet-9 PatchGAN

- **Conclusions**
  - C-GAN is effective in translating image from one domain to another.
  - Residual connection is very effective in image translation tasks because it makes deeper networks easier to train.

- **Future Work**
  - Explore residual-based network for discriminator
  - Experiment with dynamic training frequency that allows generator to train more often than discriminator in the beginning and gradually slow down.

---

**CS 230 Poster Session, Deep Learning, Spring 2018 – 9 June 2018**