Image domain translation with GANS

Hidde Lekanne gezegd Deprez
Motivation

- There are lots of simulations
- Improve them all
Acknowledgements

- Big thanks to the HULKs
- Thank you B-Human
What did I do?
Tried to close this reality gap with a GAN architecture
How did I do it?
Generative Adversarial Networks

Diagram showing the process of Generative Adversarial Networks (GANs):
- Random Noise
  - Input to Generator
  - Output: Generated Image
  - Output: Sampled Image
- Data
  - Input to Discriminator
  - Output: Discriminator Loss
- Generator
  - Input: Random Noise
  - Output: Generated Image
- Discriminator
  - Input: Generated Image, Sampled Image
  - Output: Generator Loss, Discriminator Loss

Image-to-image translation
Image-to-image translation

Some considerations:

● Unsupervised vs supervised
● One to many vs one to one
● Detangled or not
Variational Autoencoders
CycleGan
CycleGan

- Cycle loss $A \rightarrow B \rightarrow A' = A$
- Identity loss $A \rightarrow A' = A$
- Discriminator loss $A \rightarrow B = \text{Sample of } B$
Adjusted loss

- Only do the A->B->A from the simulation
- Increase the weights of cycle loss
How good did I do?
Evaluation

- Frechet Inception Distance (FID)
- Train a Instance Segmenter
FIDs

- Curvature similarity over datasets
- Datasets approximated as multivariate gaussian

<table>
<thead>
<tr>
<th></th>
<th>Simulation</th>
<th>Real</th>
</tr>
</thead>
<tbody>
<tr>
<td>CycleGan</td>
<td>207.270</td>
<td>52.773</td>
</tr>
<tr>
<td>CycleGan adjusted loss</td>
<td>176.545</td>
<td>38.085</td>
</tr>
<tr>
<td>MUNIT</td>
<td>212.587</td>
<td>38.982</td>
</tr>
<tr>
<td>MUNIT adjusted loss</td>
<td>202.637</td>
<td>28.096</td>
</tr>
</tbody>
</table>
MASK-RCNN

- Instance segmenter
- Bounding boxes first, then segmentation
- RCNN = Region convolutional neural net (not recurrent)
## MASK-RCNN

- Segmentations Average Precision scores

### Table 1: Average Precision Scores

<table>
<thead>
<tr>
<th>Method</th>
<th>lines</th>
<th>robots</th>
<th>goalframe</th>
<th>goalnet</th>
<th>field</th>
<th>ball</th>
<th>background</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pure Simulation</td>
<td>4.603</td>
<td>72.623</td>
<td>50.758</td>
<td>81.292</td>
<td>92.158</td>
<td>80.130</td>
<td>92.364</td>
</tr>
<tr>
<td>CycleGan</td>
<td>4.76</td>
<td>64.572</td>
<td>41.892</td>
<td>73.729</td>
<td>92.424</td>
<td>70.392</td>
<td>90.300</td>
</tr>
<tr>
<td>CycleGan adjusted loss</td>
<td><strong>4.88</strong></td>
<td>67.298</td>
<td>44.736</td>
<td>76.739</td>
<td>92.539</td>
<td>72.374</td>
<td>90.936</td>
</tr>
<tr>
<td>MUNIT</td>
<td>3.803</td>
<td>54.248</td>
<td>29.565</td>
<td>63.974</td>
<td>90.921</td>
<td>49.814</td>
<td>86.036</td>
</tr>
<tr>
<td>MUNIT adjusted loss</td>
<td>4.201</td>
<td>58.058</td>
<td>33.557</td>
<td>66.484</td>
<td>91.227</td>
<td>61.317</td>
<td>87.351</td>
</tr>
</tbody>
</table>

### Table 2: Number of Robots and Ball

<table>
<thead>
<tr>
<th>Method</th>
<th>robots</th>
<th>ball</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pure Simulation</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CycleGan</td>
<td>47.479</td>
<td>49.497</td>
</tr>
<tr>
<td>CycleGan adjusted loss</td>
<td>49.232</td>
<td>42.327</td>
</tr>
<tr>
<td>MUNIT</td>
<td>55.926</td>
<td>50.835</td>
</tr>
<tr>
<td>MUNIT adjusted loss</td>
<td><strong>56.705</strong></td>
<td><strong>50.910</strong></td>
</tr>
</tbody>
</table>
Panoptic Segmentation

- Fusion of semantic segmentation and object segmentation
Questions?
Bonus
Bonus