Unpaired Image-to-Image Translation with CycleGAN

Jun-Yan Zhu and Taesung Park
Joint work with Phillip Isola and Alexei A. Efros
Image-to-Image Translation with *pix2pix*

Image-to-image Translation with Conditional Adversarial Nets
Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros. CVPR 2017
Paired

\( x_i \), \( y_i \)

\{ \}

\{ \}

\{ \}

\{ \}

\{ \}

\{ \}

\{ \}

\{ \}

\{ \}

\{ \}
Paired
test.

- Expensive to collect pairs.
- Impossible in many scenarios.

Label ↔ photo: per-pixel labeling

Horse ↔ zebra: how to get zebras?
Paired

\( x_i \)

\( y_i \)

Unpaired

\( X \)

\( Y \)
No input-output pairs!
Discriminator $D$ receives $G(x)$ and determines if it is real.
GANs do **not** force output to correspond to input
Cycle-Consistent Adversarial Networks

[X] → [Y] → [X]  

[D_Y]  

[Zhu*, Park*, Isola, and Efros, ICCV 2017]
Cycle-Consistent Adversarial Networks

[Mark Twain, 1903]

[Zhu*, Park*, Isola, and Efros, ICCV 2017]
Cycle-Consistent Adversarial Networks

\[ x \xrightarrow{G} \hat{Y} \xrightarrow{F} \hat{X} \]

\[ D_Y(G(x)) \]

Reconstruction error

\[ \| F(G(x)) - x \|_1 \]

[Zhu*, Park*, Isola, and Efros, ICCV 2017]
Cycle Consistency Loss

\[ \|F(G(x)) - x\|_1 \]

[Zhu*, Park*, Isola, and Efros, ICCV 2017]
Cycle Consistency Loss

\[
\text{Reconstruction error} = \|F(G(x)) - x\|_1
\]

\[
\text{Reconstruction error} = \|G(F(y)) - y\|_1
\]

See similar formulations [Yi et al. 2017], [Kim et al. 2017], [Zhu*, Park*, Isola, and Efros, ICCV 2017]
Cycle Consistency in Vision

Consistent Track
Inconsistent Track

Forward-Backward Error: Automatic Detection of Tracking Failures. ICPR 10’
Zdenek Kalal, Krystian Mikolajczyk, and Jiri Matas.
Also see [Sundaram, Brox, Keutzer, ECCV 10’]
Cycle Consistency in Vision

Shape Matching

Co-segmentation

SfM

Huang et al, SGP’13

Wang et al, ICCV’13

Zach et al, CVPR’10

Collection Correspondence

Zhou et al, CVPR’15

Zhou et al, ICCV’15

slides credit @Tinghui Zhou
Results
<table>
<thead>
<tr>
<th>Loss</th>
<th>Map → Photo</th>
<th>Photo → Map</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% Turkers labeled real</td>
<td>% Turkers labeled real</td>
</tr>
<tr>
<td>CoGAN [30]</td>
<td>0.6% ± 0.5%</td>
<td>0.9% ± 0.5%</td>
</tr>
<tr>
<td>BiGAN/ALI [8, 6]</td>
<td>2.1% ± 1.0%</td>
<td>1.9% ± 0.9%</td>
</tr>
<tr>
<td>SimGAN [45]</td>
<td>0.7% ± 0.5%</td>
<td>2.6% ± 1.1%</td>
</tr>
<tr>
<td>Feature loss + GAN</td>
<td>1.2% ± 0.6%</td>
<td>0.3% ± 0.2%</td>
</tr>
<tr>
<td>CycleGAN (ours)</td>
<td><strong>26.8% ± 2.8%</strong></td>
<td><strong>23.2% ± 3.4%</strong></td>
</tr>
</tbody>
</table>

AMT ‘real vs fake’ test on maps ↔ aerial

<table>
<thead>
<tr>
<th>Loss</th>
<th>Per-pixel acc.</th>
<th>Per-class acc.</th>
<th>Class IOU</th>
</tr>
</thead>
<tbody>
<tr>
<td>CoGAN [30]</td>
<td>0.40</td>
<td>0.10</td>
<td>0.06</td>
</tr>
<tr>
<td>BiGAN/ALI [8, 6]</td>
<td>0.19</td>
<td>0.06</td>
<td>0.02</td>
</tr>
<tr>
<td>SimGAN [45]</td>
<td>0.20</td>
<td>0.10</td>
<td>0.04</td>
</tr>
<tr>
<td>Feature loss + GAN</td>
<td>0.06</td>
<td>0.04</td>
<td>0.01</td>
</tr>
<tr>
<td>CycleGAN (ours)</td>
<td><strong>0.52</strong></td>
<td><strong>0.17</strong></td>
<td><strong>0.11</strong></td>
</tr>
</tbody>
</table>

FCN scores on cityscapes labels → photos

<table>
<thead>
<tr>
<th>Loss</th>
<th>Per-pixel acc.</th>
<th>Per-class acc.</th>
<th>Class IOU</th>
</tr>
</thead>
<tbody>
<tr>
<td>CoGAN [30]</td>
<td>0.45</td>
<td>0.11</td>
<td>0.08</td>
</tr>
<tr>
<td>BiGAN/ALI [8, 6]</td>
<td>0.41</td>
<td>0.13</td>
<td>0.07</td>
</tr>
<tr>
<td>SimGAN [45]</td>
<td>0.47</td>
<td>0.11</td>
<td>0.07</td>
</tr>
<tr>
<td>Feature loss + GAN</td>
<td>0.50</td>
<td>0.10</td>
<td>0.06</td>
</tr>
<tr>
<td>CycleGAN (ours)</td>
<td><strong>0.58</strong></td>
<td><strong>0.22</strong></td>
<td><strong>0.16</strong></td>
</tr>
</tbody>
</table>

Classification performance of photo → labels
Collection Style Transfer

Photograph @ Alexei Efros

Monet

Van Gogh

Cezanne

Ukiyo-e
Monet’s paintings → photos
Monet’s paintings → photos
Why CycleGAN works
Style and Content Separation

**Paired Separation**

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>?</th>
<th>?</th>
<th>?</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Unpaired Separation**

Adversarial Loss: change the style

\[
\mathcal{L}_{GAN}(G, D_Y, X, Y) = E_{y \sim p_{data}(y)} [\log D_Y(y)] + E_{x \sim p_{data}(x)} [\log(1 - D_Y(G(x))]
\]

Cycle Consistency Loss: preserve the content

\[
\mathcal{L}_{cyc}(G, F) = E_{x \sim p_{data}(x)} [\|F(G(x)) - x\|_1] + E_{y \sim p_{data}(y)} [\|G(F(y)) - y\|_1].
\]

Two empirical assumptions:
- content is easy to keep.
- style is easy to change.
Neural Style Transfer [Gatys et al. 2015]

Style and Content:
- Content: feature difference
- Style: Gram Matrix difference
- Both losses are hard-coded.
Photo → Van Gogh

horse → zebra
Cycle Loss \textbf{upper bounds} Conditional Entropy

“ALICE: Towards Understanding Adversarial Learning for Joint Distribution Matching” [Li et al. NIPS 2017]. Also see [Tiao et al. 2018] “CycleGAN as Approximate Bayesian Inference”

Conditional Entropy

\[ H^\pi(x|z) \triangleq -\mathbb{E}_{\pi(x,z)}[\log \pi(x|z)] \]
Cycle Loss upper bounds Conditional Entropy

**Lemma 3** For joint distributions $p_\theta(x, z)$ or $q_\phi(x, z)$, we have

$$H^{q_\phi}(x|z) \triangleq -\mathbb{E}_{q_\phi(x,z)}[\log q_\phi(x|z)] = -\mathbb{E}_{q_\phi(x,z)}[\log p_\theta(x|z)] - \mathbb{E}_{q_\phi(z)}[\text{KL}(q_\phi(x|z)||p_\theta(x|z))]$$

$$\leq -\mathbb{E}_{q_\phi(x,z)}[\log p_\theta(x|z)] \triangleq \mathcal{L}_{\text{Cycle}}(\theta, \phi).$$

“ALICE: Towards Understanding Adversarial Learning for Joint Distribution Matching” [Li et al. NIPS 2017]. Also see [Tiao et al. 2018] “CycleGAN as Approximate Bayesian Inference”
CycleGAN implementations

Torch

- pytorch-CycleGAN-and-pix2pix
  Image-to-image translation in PyTorch (e.g., horse2zebra, edges2cats, and more)
  - Python
  - 4.3k stars
  - 970 forks

PyTorch

- CycleGAN
  Software that can generate photos from paintings, turn horses into zebras, perform style transfer, and more.
  - Lua
  - 6.5k stars
  - 940 forks

20+ implementations by researchers/developers:
- Tensorflow, Chainer, mxnet, Lasagne, Keras...
CycleGAN at School

- Taught at Stanford, UC Berkeley, UoT, Udacity, FastAI, etc.
- Course assignment code and handout designed by Prof. Roger Grosse for CSC321 “Intro to Neural Networks and Machine Learning” at University of Toronto.

```python
## FILL THIS IN: CREATE ARCHITECTURE ##

# 1. Define the encoder part of the generator
# self.conv1 = ...
# self.conv2 = ...

# 2. Define the transformation part of the generator
# self.resnet_block = ...

# 3. Define the decoder part of the generator
# self.deconv1 = ...
# self.deconv2 = ...
```
Applications
CG2Real: GTA5 $\rightarrow$ real streetview

Inspired by [Johnson et al. 2011]
Real2CG: real streetview → GTA

Cityscape Input

Output
Synthetic Data as Supervision

GTA5 images

Segmentation labels

[Richter*, Vineet* et al. 2016] [Krähenbühl et al. 2018]
## Domain Adaptation with CycleGAN

**Train on GTA5 data**

<table>
<thead>
<tr>
<th></th>
<th>meanIOU</th>
<th>Per-pixel accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oracle (Train and test on Real)</td>
<td>60.3</td>
<td>93.1</td>
</tr>
<tr>
<td>Train on CG, test on Real</td>
<td>17.9</td>
<td>54.0</td>
</tr>
</tbody>
</table>

**Test on real images**

See Judy Hoffman’s talk at 14:30 “Adversarial Domain Adaptation”
Domain Adaptation with CycleGAN

GTA5 data + Domain adaptation

<table>
<thead>
<tr>
<th></th>
<th>meanIOU</th>
<th>Per-pixel accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oracle (Train and test on Real)</td>
<td>60.3</td>
<td>93.1</td>
</tr>
<tr>
<td>Train on CG, test on Real</td>
<td>17.9</td>
<td>54.0</td>
</tr>
<tr>
<td>FCN in the wild [Previous STOA]</td>
<td>27.1</td>
<td>-</td>
</tr>
</tbody>
</table>

Test on real images

See Judy Hoffman’s talk at 14:30 “Adversarial Domain Adaptation”
Domain Adaptation with CycleGAN

<table>
<thead>
<tr>
<th></th>
<th>meanIOU</th>
<th>Per-pixel accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oracle (Train and test on Real)</td>
<td>60.3</td>
<td>93.1</td>
</tr>
<tr>
<td>Train on CG, test on Real</td>
<td>17.9</td>
<td>54.0</td>
</tr>
<tr>
<td>FCN in the wild [Previous STOA]</td>
<td>27.1</td>
<td>-</td>
</tr>
<tr>
<td>Train on CycleGAN, test on Real</td>
<td>34.8</td>
<td>82.8</td>
</tr>
</tbody>
</table>

See Judy Hoffman’s talk at 14:30 “Adversarial Domain Adaptation”
Applications and Extensions

Attribute Editing [Lu et al.]
- Low-res
- Bald
- Bangs
  - arXiv:1705.09966

Object Editing [Liang et al.]
- Mask
- Input
- Output
  - arXiv:1708.00315

Front/Character Transfer [Ignatov et al.]
- Input
- output
  - arXiv:1801.08624

Data generation [Wang et al.]
- samples by CycleWGAN
  - arXiv:1707.03124
Photo Enhancement

WESPE: Weakly Supervised Photo Enhancer for Digital Cameras. arxiv 1709.01118
Andrey Ignatov, Nikolay Kobyshev, Kenneth Vanhoey, Radu Timofte, Luc Van Gool
Image Dehazing

Cycle-Dehaze: Enhanced CycleGAN for Single Image Dehazing. CVPRW 2018
Deniz Engin*, Anıl Genc*, Hazım Kemal Ekenel
Unsupervised Motion Retargeting

Neural Kinematic Networks for Unsupervised Motion Retargetting. CVPR 2018 (oral)
Ruben Villegas, Jimei Yang, Duygu Ceylan, Honglak Lee
Neural Kinematic Networks for Unsupervised Motion Retargetting. CVPR 2018 (oral)
Ruben Villegas, Jimei Yang, Duygu Ceylan, Honglak Lee
Applications Beyond Computer Vision

- Medical Imaging and Biology [Wolterink et al., 2017]
- Voice conversion [Fang et al., 2018, Kaneko et al., 2017]
- Cryptography [CipherGAN: Gomez et al., ICLR 2018]
- Robotics
- NLP: Text style transfer.

...
Deep MR to CT Synthesis using Unpaired Data

Jelmer M. Wolterink\textsuperscript{1,}\textsuperscript{2}, Anna M. Dinkla\textsuperscript{2}, Mark H.F. Savenije\textsuperscript{2}, Peter R. Seevinck\textsuperscript{1}, Cornelis A.T. van den Berg\textsuperscript{2}, Ivana Išgum\textsuperscript{1}

\textsuperscript{1} Image Sciences Institute, University Medical Center Utrecht, The Netherlands 
\texttt{j.m.wolterink@umcutrecht.nl}

\textsuperscript{2} Department of Radiotherapy, University Medical Center Utrecht, The Netherlands
Problem with the paired MR-CT data

• Images are not perfectly aligned
• There are usually more unpaired data
Reconstucted Images
Reconstructed Images
Reconstructed Images (Chu, Zhmoginov and Sandler, NIPSW 2017)
Reconstructed Images (Chu, Zhmoginov and Sandler, NIPS 2017)
Failure cases
ImageNet
“Wild horse”
Beyond CycleGAN

Multi-modality

Style control

More than two domains
Beyond CycleGAN

Multi-modality

Style control

More than two domains
Beyond CycleGAN – Multi-modality

Augmented CycleGAN: Learning Many-to-Many Mappings from Unpaired Data

Amjad Almahairi 1†, Sai Rajeswar 1, Alessandro Sordoni 2, Philip Bachman 2, Aaron Courville 1, 3

(a) CycleGAN
(b) Augmented CycleGAN

Almahairi et al., ICML 2018
Beyond CycleGAN – Multi-modality

\[ z_2 \]

\[ a \]

\[ b \]

\[ \tilde{b} \]

\[ a' \]

\[ \tilde{a} \]

\[ b' \]

\[ G_{AB} \]

\[ G_{BA} \]

\[ E_A \]

\[ E_B \]

Almahairi et al., ICML 2018
Beyond CycleGAN – Multi-modality

Almahairi et al., ICML 2018
Beyond CycleGAN

Multi-modality

*Style control*

More than two domains
Beyond CycleGAN – Style Control

PairedCycleGAN: Asymmetric Style Transfer for Applying and Removing Makeup

Huiwen Chang
Princeton University

Jingwan Lu
Adobe Research

Fisher Yu
UC Berkeley

Adam Finkelstein
Princeton University

Chang et al., CVPR 2018
Beyond CycleGAN – Style Control

Chang et al., CVPR 2018
Beyond CycleGAN – Style Control

Source | Reference | Our result

Chang et al., CVPR 2018
Beyond CycleGAN

Multi-modality

Style control

More than two domains
Beyond CycleGAN – More than 2 domains

StarGAN: Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation

Yunjey Choi\textsuperscript{1,2}, Minje Choi\textsuperscript{1,2}, Munyoung Kim\textsuperscript{2,3}, Jung-Woo Ha\textsuperscript{2}, Sunghun Kim\textsuperscript{2,4}, Jaegul Choo\textsuperscript{1,2}

\textsuperscript{1} Korea University \quad \textsuperscript{2} Clova AI Research, NAVER \quad \textsuperscript{3} The College of New Jersey \quad \textsuperscript{4} Hong Kong University of Science & Technology
Beyond CycleGAN – More than 2 domains

Choi et al., CVPR 2018
Beyond CycleGAN – More than 2 domains

Choi et al., CVPR 2018
Beyond CycleGAN – More than 2 domains
Beyond CycleGAN – More than 2 domains

Anoosheh et al., CVPR 2018
#CycleGAN at Twitter

Monet → Thomas Kinkade @David Fouhey

Resurrecting Ancient Cities @ Jack Clark

Birds @Matt Powell

Bear → Panda @Matt Powell
Turn Real People Into Anime Art

Results

Ongoing work
Still improving it

@minjunli (Minjun Li), @Aixile (Yanghua JIN), @alanyttian (Yingtao Tian)
Latest from #CycleGAN

CycleGAN with architectural modifications, by itok_msi
https://qiita.com/itok_msi/items/b6b615bc28b1a720afd7
Thank You!

Code and data: junyanz.github.io/CycleGAN/