Unsupervised Conditional Generation
Unsupervised Conditional Generation

Transform an object from one domain to another **without paired data** (e.g. style transfer)
Unsupervised Conditional Generation

• Approach 1: Direct Transformation

\[ G_{X \rightarrow Y} \]

Domain X \[ \rightarrow \] \[ G_{X \rightarrow Y} \] \[ \rightarrow \] Domain Y

For texture or color change

• Approach 2: Projection to Common Space

\[ EN_X \]

Encoder of domain X \[ \rightarrow \] \[ DE_Y \]

Face Attribute

Decoder of domain Y

Domain X \[ \rightarrow \] \[ EN_X \] \[ \rightarrow \] \[ DE_Y \] \[ \rightarrow \] Domain Y

Larger change, only keep the semantics
Direct Transformation

\[ G_{X \to Y} \]

Become similar to domain Y

Input image belongs to domain Y or not

\[ D_Y \]

scalar
Direct Transformation

Domain X

\[ G_{X \rightarrow Y} \]

Become similar to domain Y

Not what we want!

\[ D_Y \]

scalar

Input image belongs to domain Y or not

ignore input

Domain Y
Direct Transformation

Domain X

\[ G_{X \rightarrow Y} \]

Become similar to domain Y

Not what we want!

Domain Y

\[ D_Y \rightarrow \text{scalar} \]

Input image belongs to domain Y or not

ignore input

The issue can be avoided by network design.
Simpler generator makes the input and output more closely related.

[Tomer Galanti, et al. ICLR, 2018]
Direct Transformation

\[ G_{X \rightarrow Y} \]

Become similar to domain Y

Encoder Network

pre-trained

as close as possible

Encoder Network

Input image belongs to domain Y or not

Baseline of DTN [Yaniv Taigman, et al., ICLR, 2017]
Direct Transformation

\[ G_{X \rightarrow Y} \rightarrow G_{Y \rightarrow X} \]

as close as possible

Cycle consistency

Lack of information for reconstruction

Input image belongs to domain Y or not

Domain Y

\[ D_{Y} \rightarrow \text{scalar} \]
Direct Transformation

as close as possible

scalar: belongs to domain X or not

as close as possible

scalar: belongs to domain Y or not
Cycle GAN – Silver Hair

- https://github.com/Aixile/chainer-cyclegan
Cycle GAN – Silver Hair

- https://github.com/Aixile/chainer-cyclegan
Issue of Cycle Consistency

• CycleGAN: a Master of Steganography (隱寫術)

\[ G_{X \rightarrow Y} \rightarrow G_{Y \rightarrow X} \]

The information is hidden.

[Casey Chu, et al., NIPS workshop, 2017]
**Disco GAN**


**Dual GAN**

[Taeksoo Kim, et al., ICML, 2017]

**Cycle GAN**

[Zili Yi, et al., ICCV, 2017]

StarGAN

For multiple domains, considering starGAN

[Yunjey Choi, arXiv, 2017]
StarGAN
StarGAN

(a) Training the discriminator
(1) when training with real images
(2) when training with fake images

(b) Original-to-target domain
Output image and original domain label

(c) Target-to-original domain
Reconstructed image

(d) Fooling the discriminator
Real?
Input image and target domain label
StarGAN

**Training the discriminator**

- **Real image**
- **Fake image**

- **CelebA label**
  - Black / Blond / Brown / Male / Young

- **RaFD label**
  - Angry / Fearful / Happy / Sad / Disgusted

- **Mask vector**
  - CelebA / RaFD

**Training with RaFD**

- (e) Training the discriminator
- (f) Original-to-target domain
- (g) Target-to-original domain
- (h) Fooling the discriminator

- **Output image and original domain label**

- **Input image and target domain label**

- **Reconstructed image**

(1) when training with real images
(2) when training with fake images
Unsupervised Conditional Generation

• Approach 1: Direct Transformation

\[ G_{X \rightarrow Y} \]

Domain X \rightarrow Domain Y

For texture or color change

• Approach 2: Projection to Common Space

\[ EN_X \rightarrow DE_Y \]

Domain X \rightarrow Domain Y

Encoder of domain X \rightarrow Face Attribute \rightarrow Decoder of domain Y

Larger change, only keep the semantics
Projection to Common Space

Target

Domain X

\[ EN_X \]

\[ DE_X \]

Domain Y

\[ EN_Y \]

\[ DE_Y \]

Face Attribute

image
Projection to Common Space

Minimizing reconstruction error

Training

Domain X

Domain Y
Projection to Common Space

Training

Minimizing reconstruction error

Because we train two auto-encoders separately ... 

The images with the same attribute may not project to the same position in the latent space.
Projection to Common Space

Training

Sharing the parameters of encoders and decoders

Couple GAN [Ming-Yu Liu, et al., NIPS, 2016]
UNIT [Ming-Yu Liu, et al., NIPS, 2017]
Minimizing reconstruction error

The domain discriminator forces the output of $E_N^X$ and $E_N^Y$ have the same distribution.

$E_N^X$ and $E_N^Y$ fool the domain discriminator

The domain discriminator forces the output of $E_N^X$ and $E_N^Y$ have the same distribution. [Guillaume Lample, et al., NIPS, 2017]
Projection to Common Space

Training

Minimizing reconstruction error

Discriminator of X domain

Discriminator of Y domain

Cycle Consistency:

Used in ComboGAN [Asha Anoosheh, et al., arXiv, 017]
Projection to Common Space

Training

To the same latent space

![Diagram](image)

Discriminator of X domain

Discriminator of Y domain

Semantic Consistency:

世界二次元化

• Using the code: https://github.com/Hi-king/kawaiii_creator
• It is not cycle GAN, Disco GAN

input

output domain
Voice Conversion
In the past

Speaker A  Speaker B

How are you?  How are you?

Good morning  Good morning

Today

Speaker A  Speaker B

天気真好  How are you?

再見囉  Good morning

Speakers A and B are talking about completely different things.
我

感謝周儒杰同學提供實驗結果
Reference


• Zili Yi, Hao Zhang, Ping Tan, Minglun Gong, DualGAN: Unsupervised Dual Learning for Image-to-Image Translation, ICCV, 2017

• Tomer Galanti, Lior Wolf, Sagie Benaim, The Role of Minimal Complexity Functions in Unsupervised Learning of Semantic Mappings, ICLR, 2018

• Yaniv Taigman, Adam Polyak, Lior Wolf, Unsupervised Cross-Domain Image Generation, ICLR, 2017

• Asha Anoosheh, Eirikur Agustsson, Radu Timofte, Luc Van Gool, ComboGAN: Unrestrained Scalability for Image Domain Translation, arXiv, 2017

• Amélie Royer, Konstantinos Bousmalis, Stephan Gouws, Fred Bertsch, Inbar Mosseri, Forrester Cole, Kevin Murphy, XGAN: Unsupervised Image-to-Image Translation for Many-to-Many Mappings, arXiv, 2017
Reference

• Guillaume Lample, Neil Zeghidour, Nicolas Usunier, Antoine Bordes, Ludovic Denoyer, Marc'Aurelio Ranzato, Fader Networks: Manipulating Images by Sliding Attributes, NIPS, 2017

• Taeksoo Kim, Moonsu Cha, Hyunsoo Kim, Jung Kwon Lee, Jiwon Kim, Learning to Discover Cross-Domain Relations with Generative Adversarial Networks, ICML, 2017

• Ming-Yu Liu, Oncel Tuzel, “Coupled Generative Adversarial Networks”, NIPS, 2016

• Ming-Yu Liu, Thomas Breuel, Jan Kautz, Unsupervised Image-to-Image Translation Networks, NIPS, 2017

• Yunjey Choi, Minje Choi, Munyoung Kim, Jung-Woo Ha, Sunghun Kim, Jaegul Choo, StarGAN: Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation, arXiv, 2017