# Unsupervised Conditional Generation

# ***Unsupervised Conditional Generation***

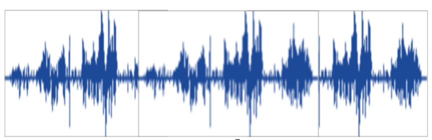
真實的畫

梵谷的畫

Domain X  → G →  Domain Y

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

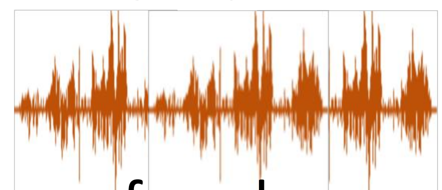
Domain X



male

↔

Domain Y



female

**Part II**

| Domain X | | Domain Y |
|---|---|---|
| It is good. | ↔ | It is bad. |
| It's a good day. | | It's a bad day. |
| I love you. | | I don't love you. |

**Part III**

# Unsupervised Conditional Generation

- Approach 1: Direct Transformation

直接轉，因此input/output不能差太多

Domain X

$G_{X \to Y}$

Domain Y

For texture or color change

- Approach 2: Projection to Common Space

利用encoder找feature，比如說眼鏡黑頭髮… 適用於input/output差比較多

Domain X

$EN_X$

Encoder of domain X

Face Attribute

$DE_Y$

Decoder of domain Y

Domain Y

Larger change, only keep the semantics

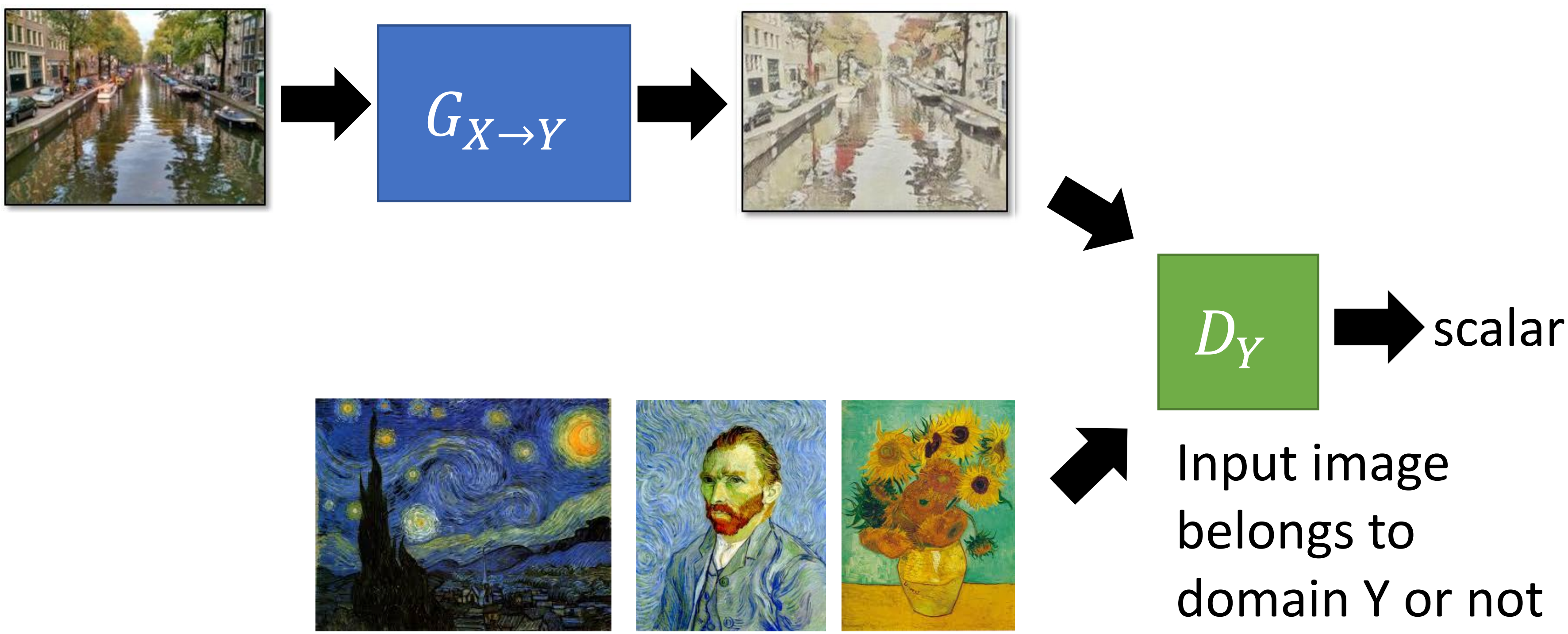
# Direct Transformation

Domain X

Domain Y

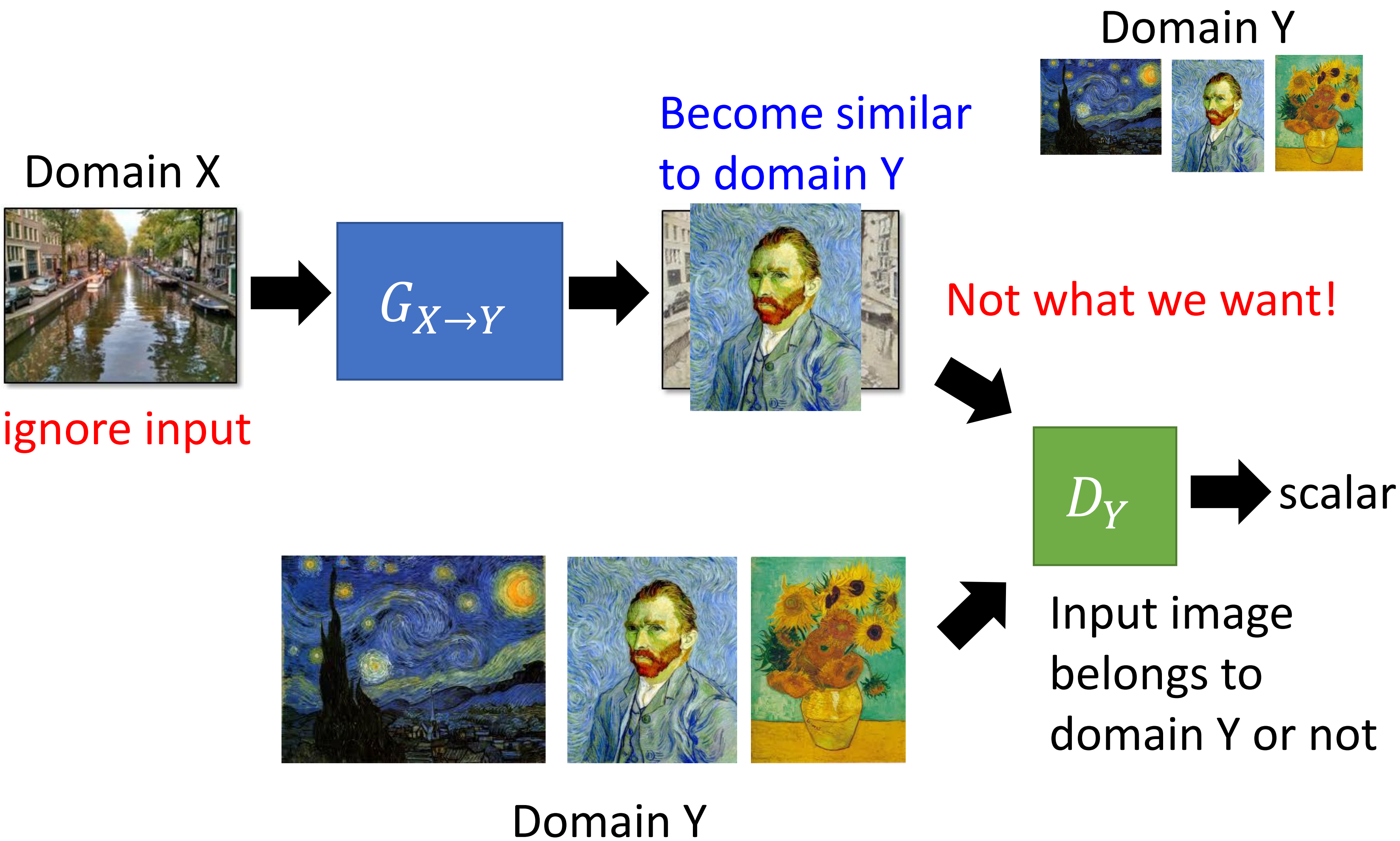
其實不加入cycle consistent也是可以work的
因為generator如果沒有很deep的話，他並不想讓
input變化太大，希望只透過一點變化即可騙過D

Domain X




Domain Y

# Direct Transformation

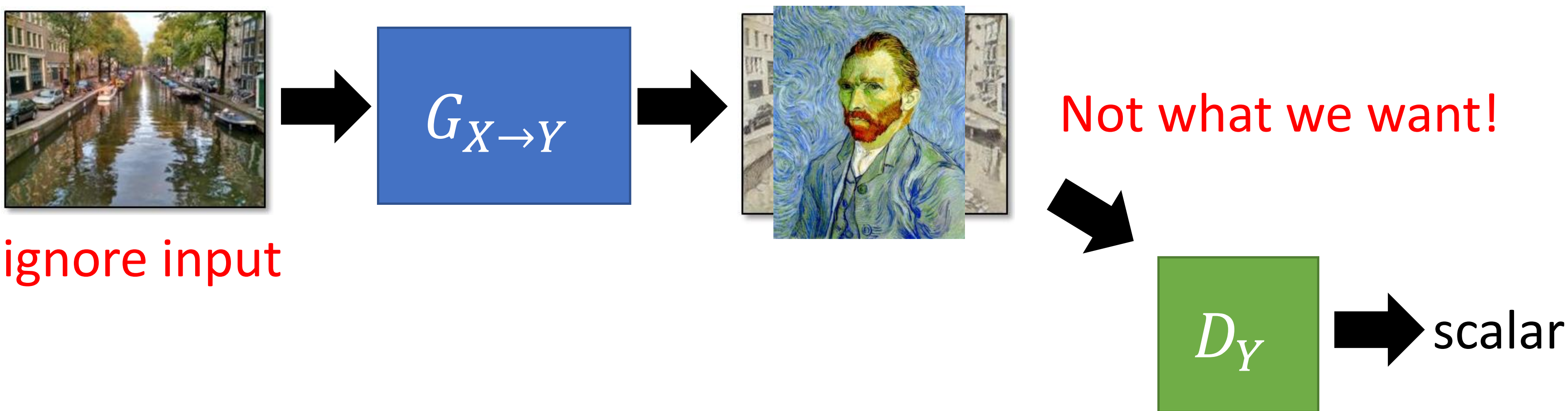
# Direct Transformation

Domain X

Domain Y

Domain X

Become similar to domain Y



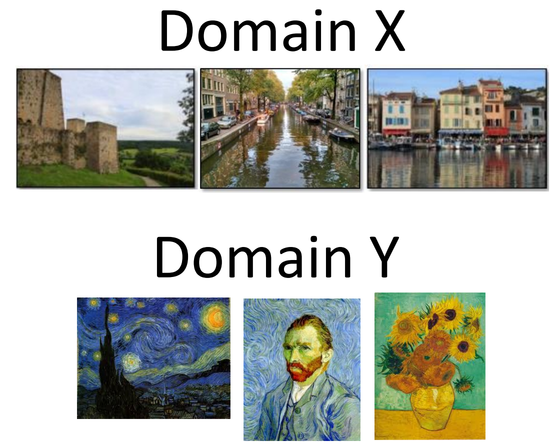

The issue can be avoided by network design.

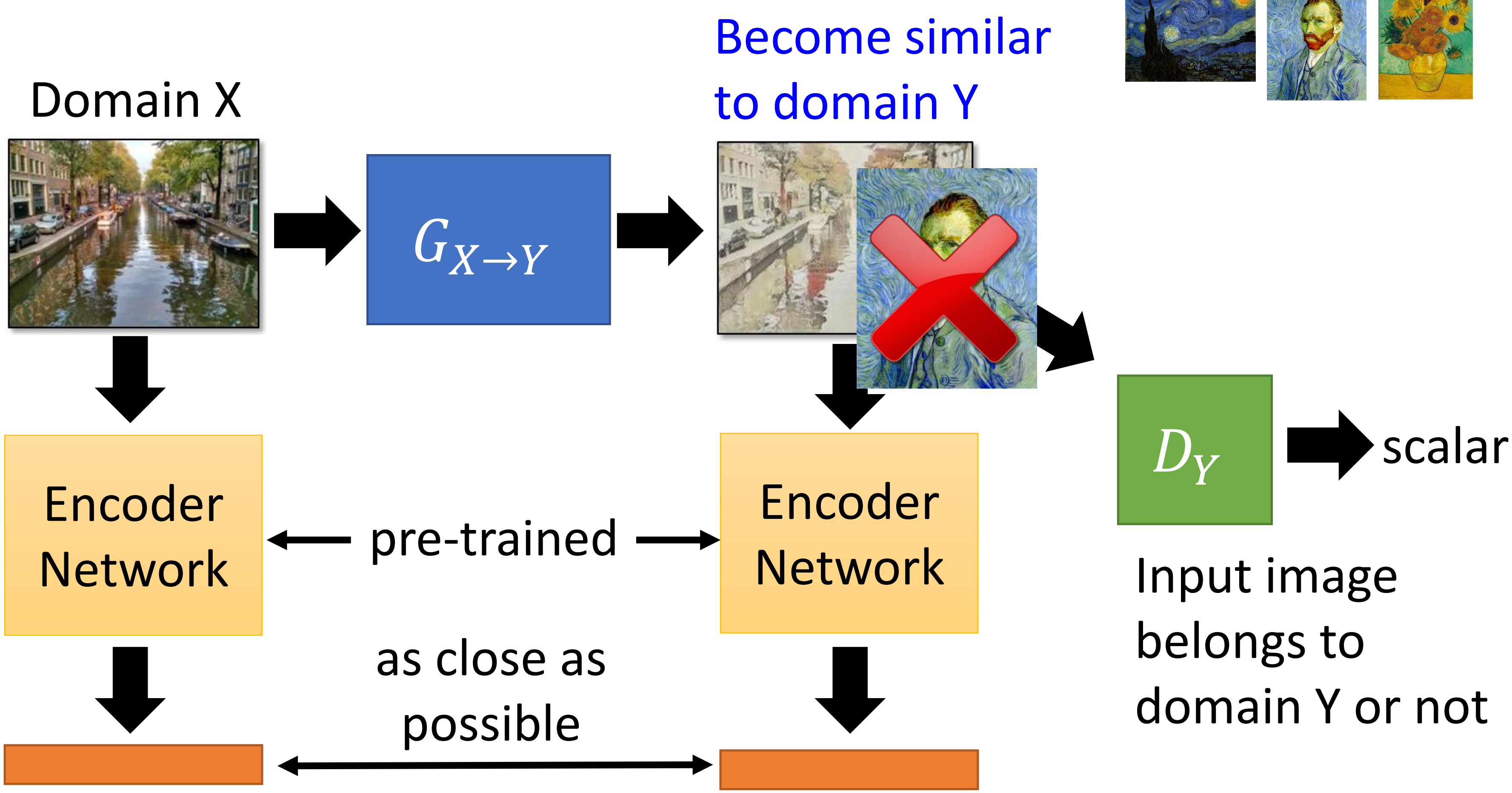
Simpler generator makes the input and output more closely related.

Input image belongs to domain Y or not

[Tomer Galanti, et al. ICLR, 2018]

# Direct Transformation

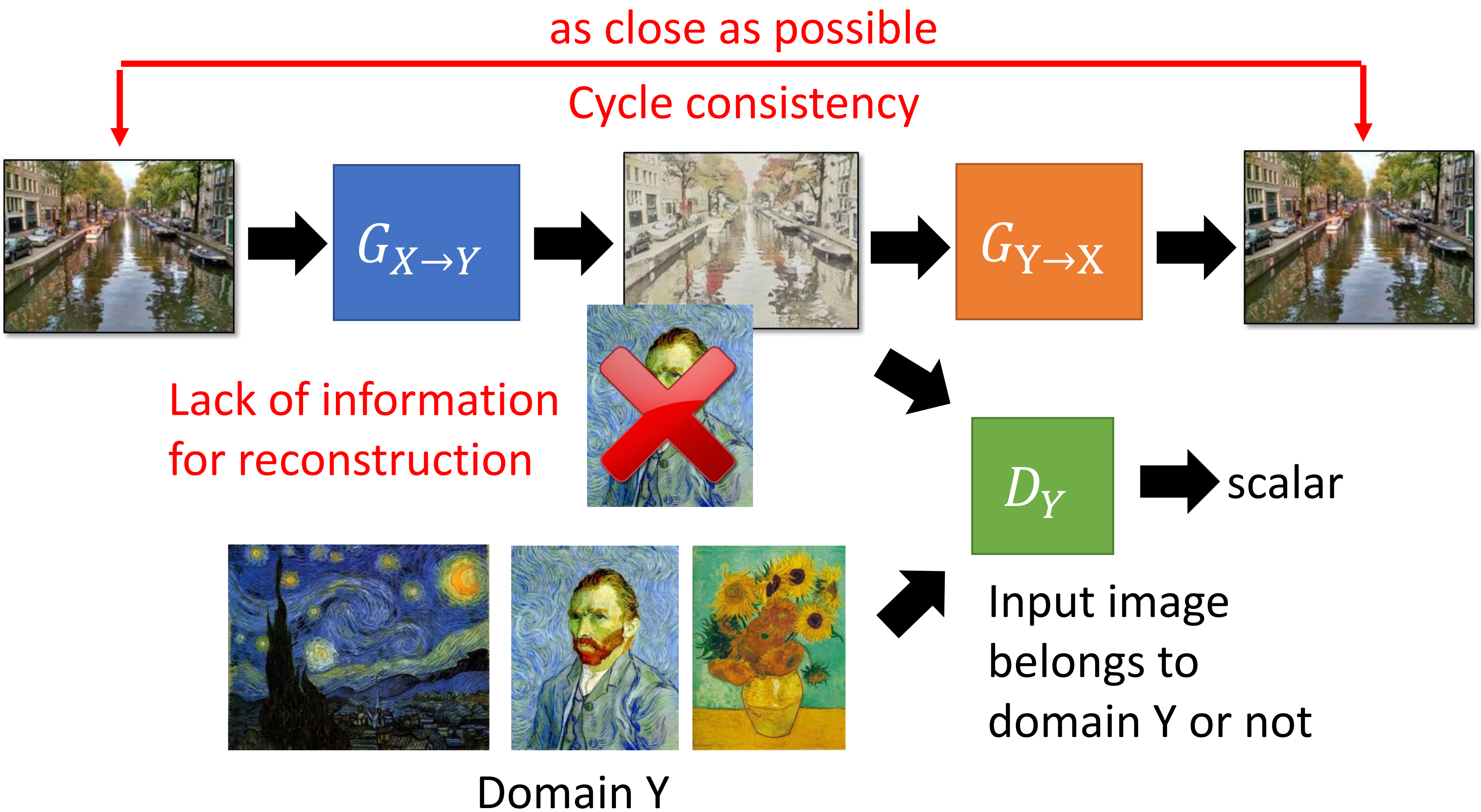


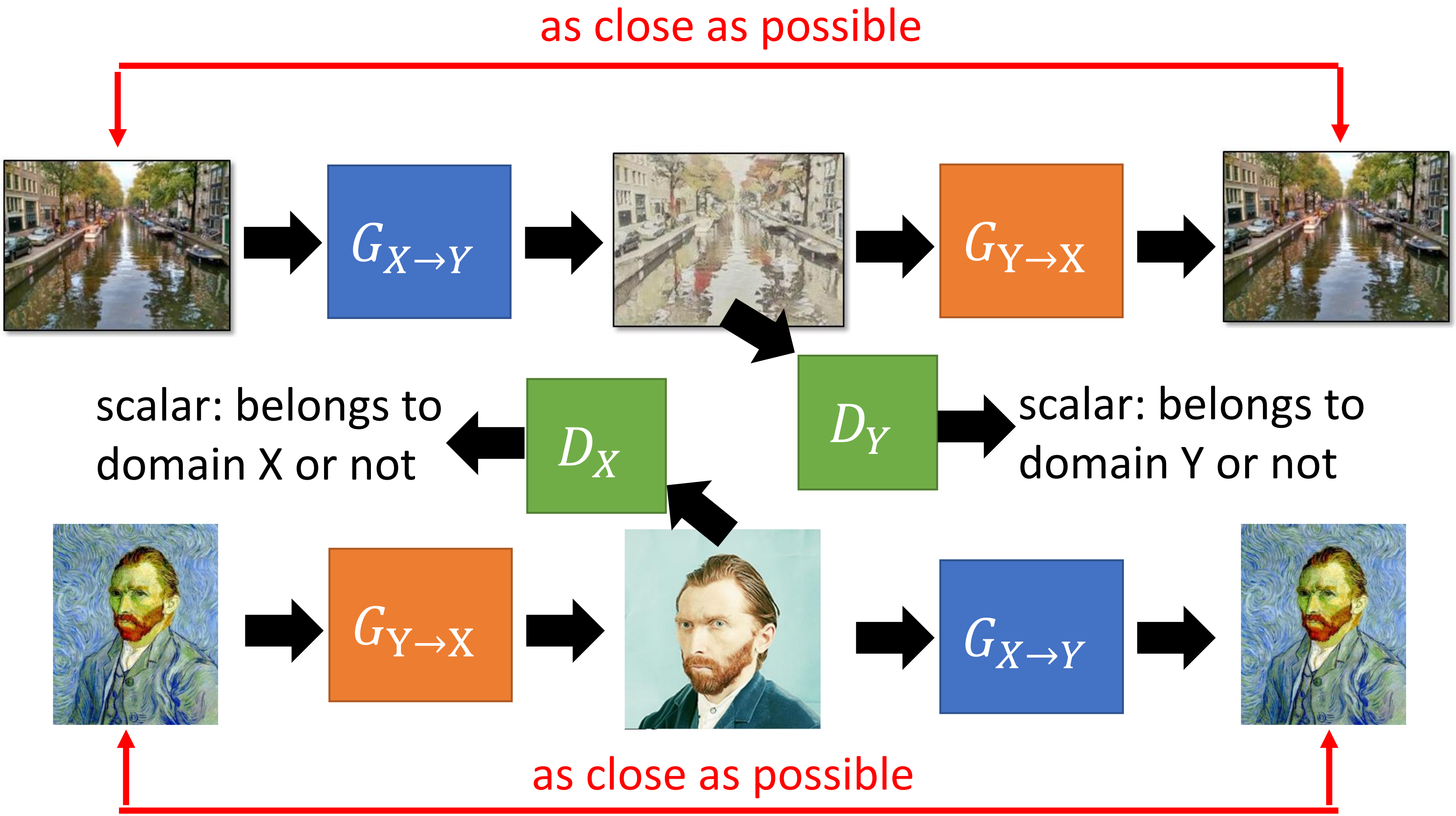

Baseline of DTN [Yaniv Taigman, et al., ICLR, 2017]

[Jun-Yan Zhu, et al., ICCV, 2017]

# Direct Transformation

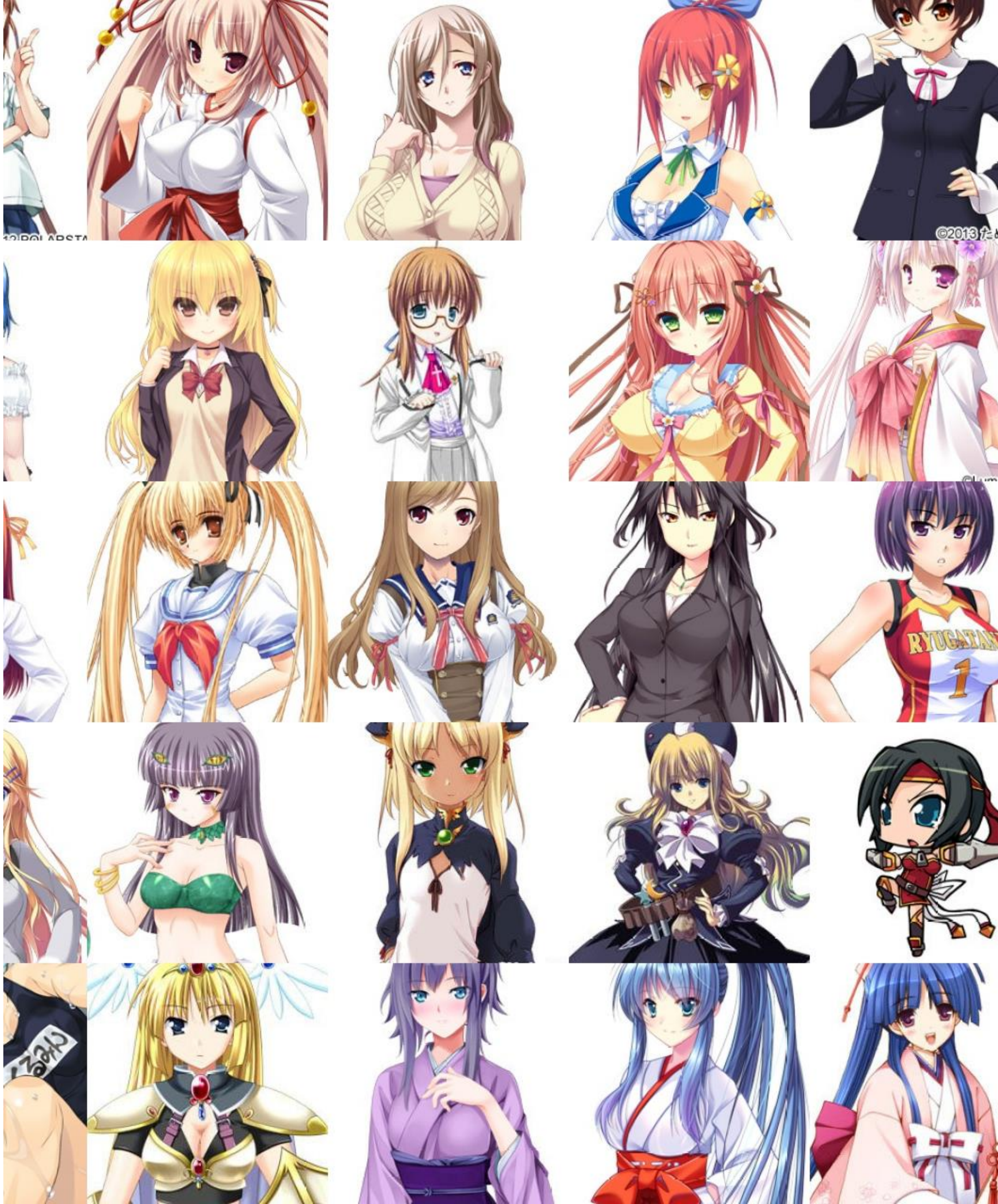


# Direct Transformation



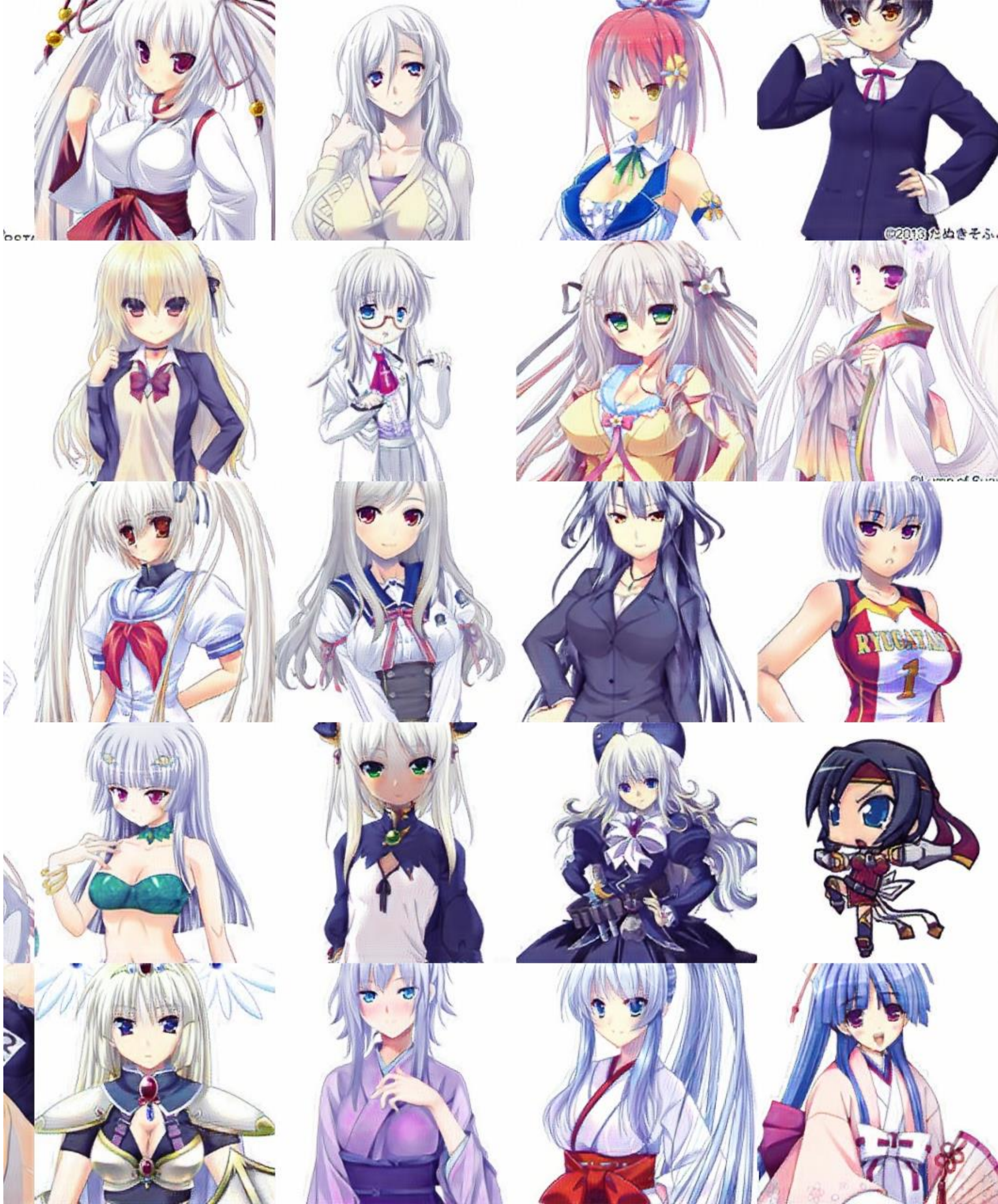
# Cycle GAN – Silver Hair

- https://github.com/Aixile/chainer-cyclegan



# Cycle GAN – Silver Hair

- https://github.com/Aixile/chainer-cyclegan



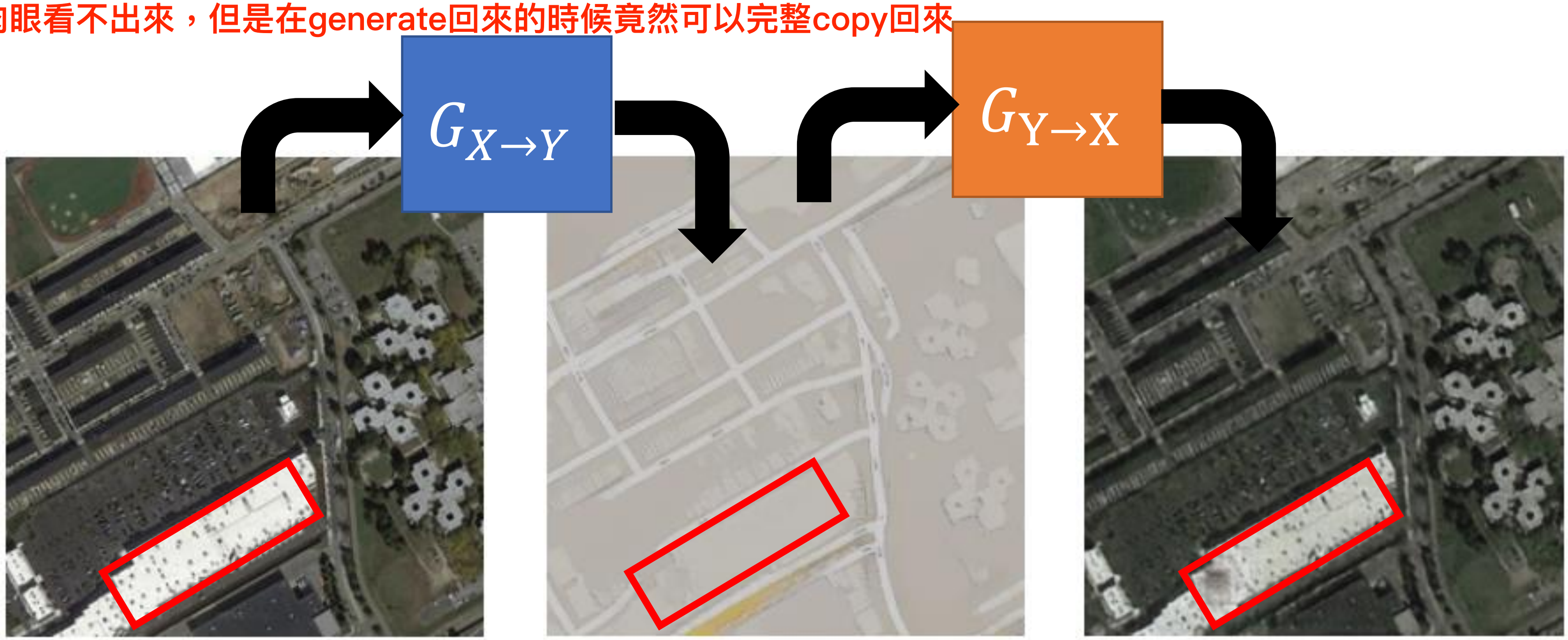
# Issue of Cycle Consistency

- **CycleGAN: a Master of Steganography** 隱寫術

cycleGAN會把input image infomation藏在output image

[Casey Chu, et al., NIPS workshop, 2017]

肉眼看不出來，但是在generate回來的時候竟然可以完整copy回來

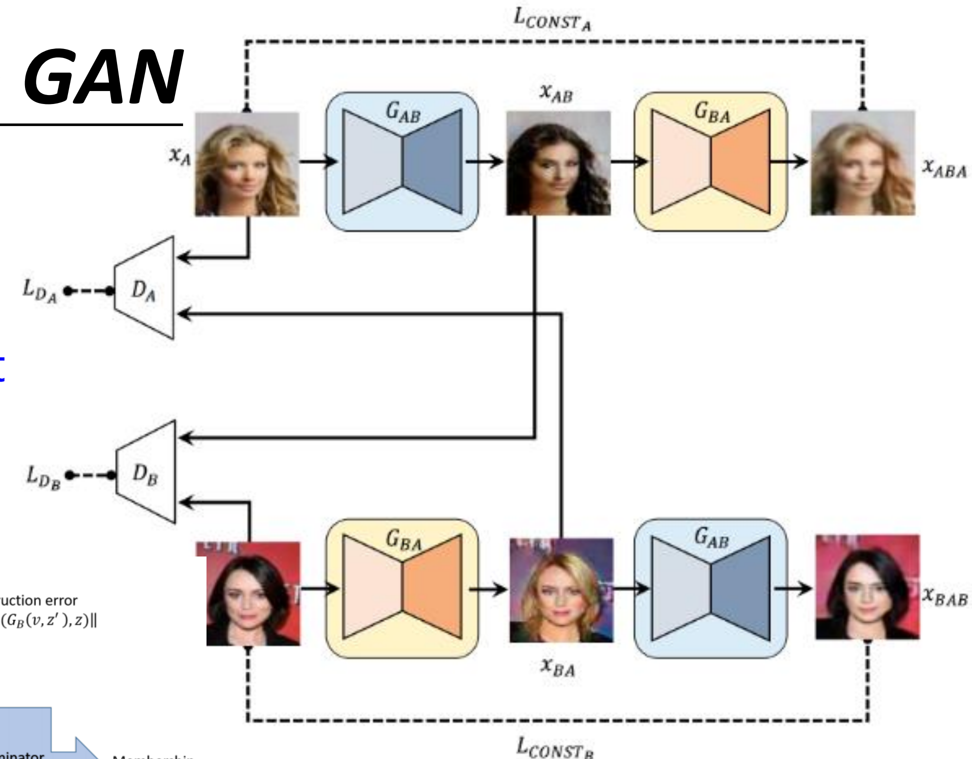



The information is hidden.

表示machine找了一個方法對付cycle consistent loss，並且不被D察覺

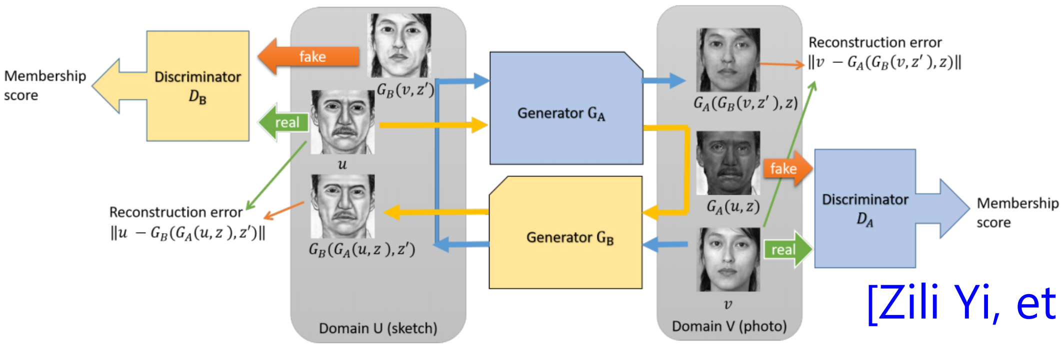
For multiple domains, considering starGAN
[Yunjey Choi, arXiv, 2017]

# Disco GAN



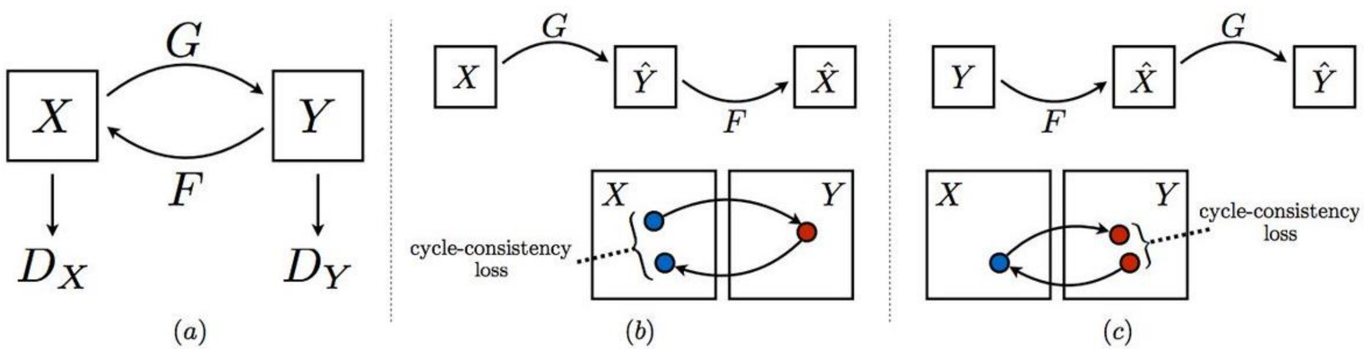

[Taeksoo Kim, et al., ICML, 2017]

# Dual GAN




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

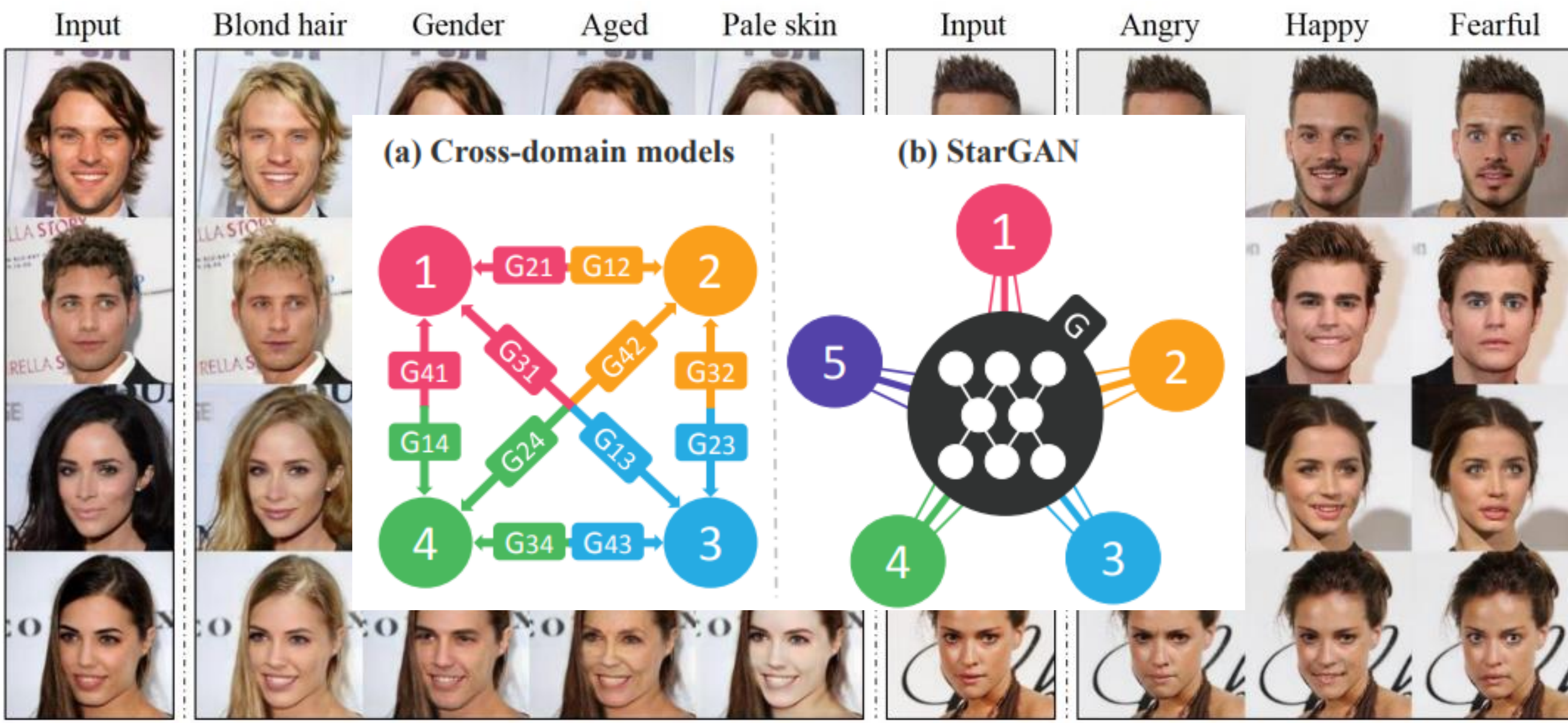
# Cycle GAN



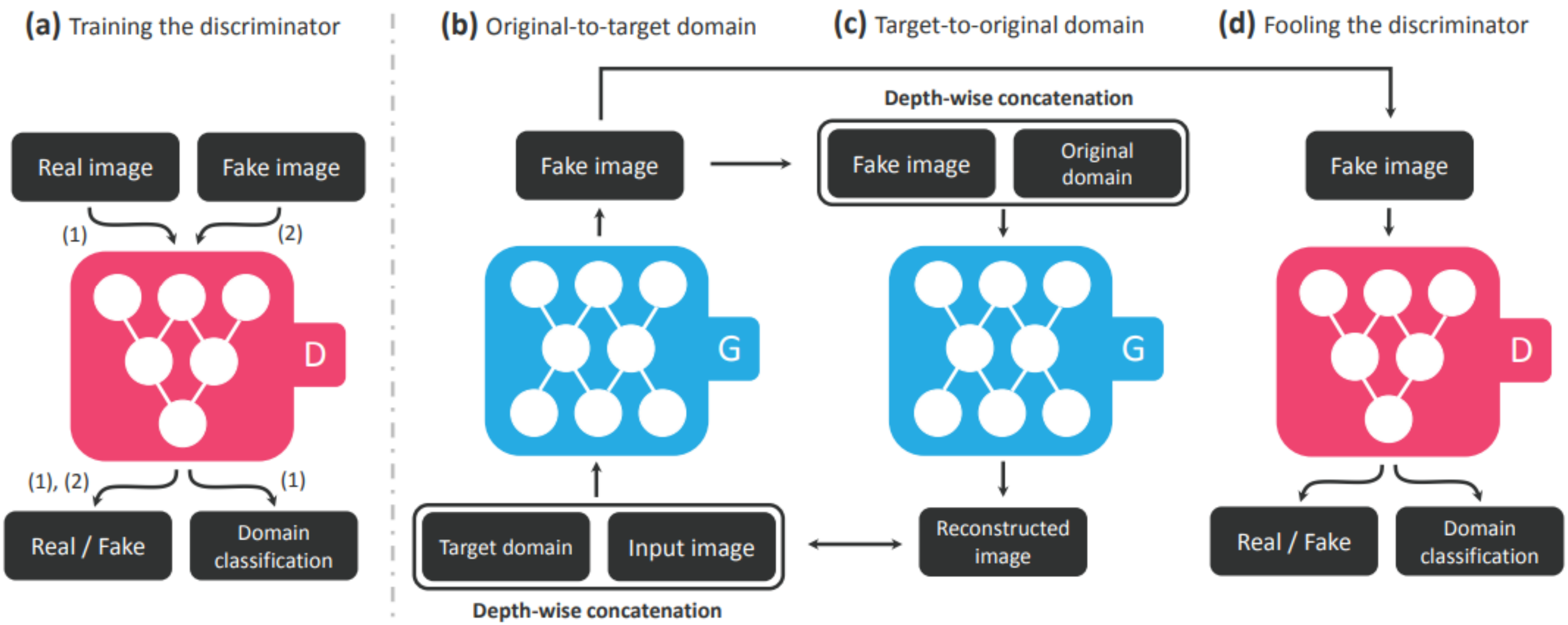

[Jun-Yan Zhu, et al., ICCV, 2017]

https://arxiv.org/pdf/1711.09020.pdf

# StarGAN

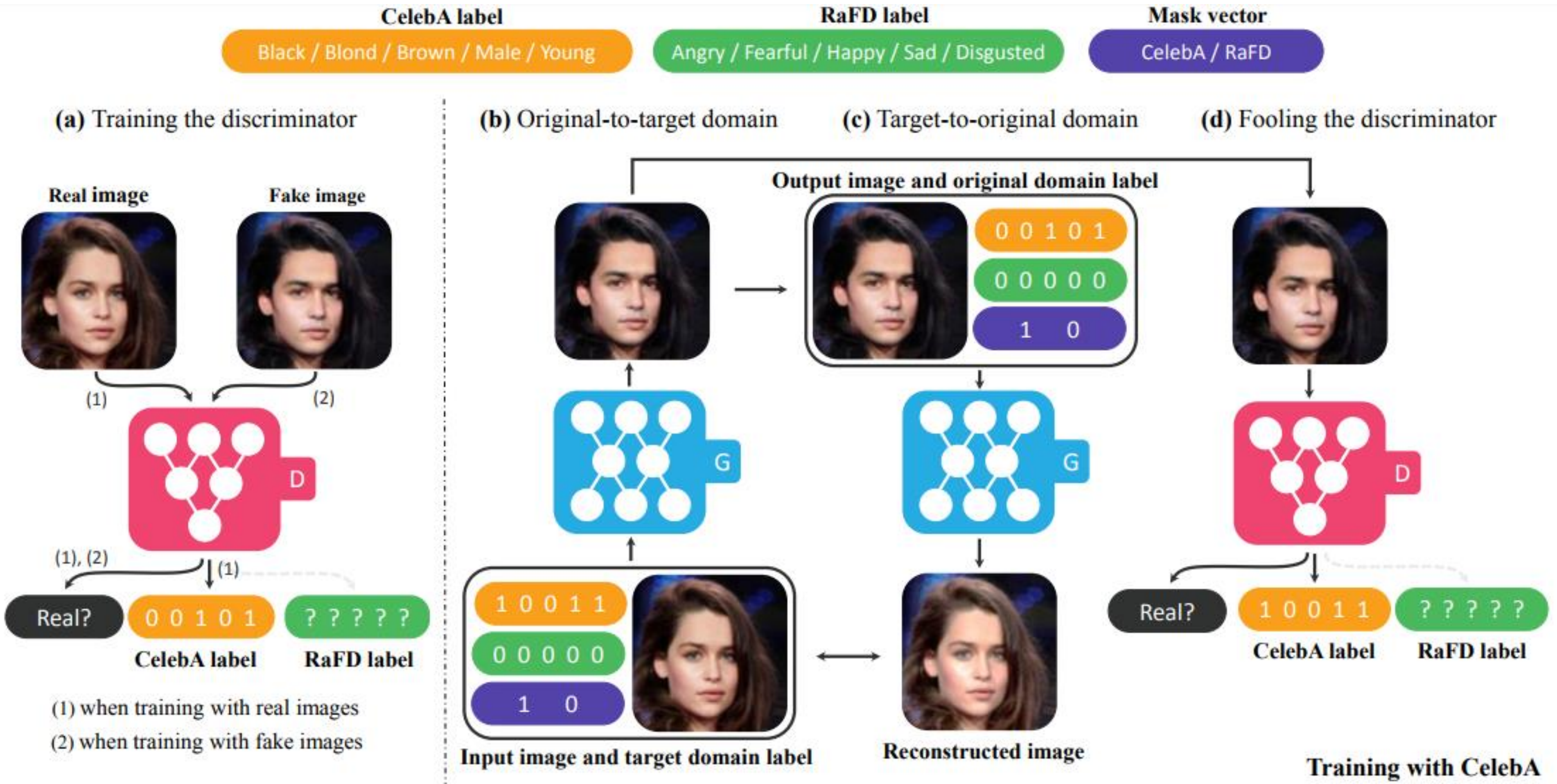


# StarGAN



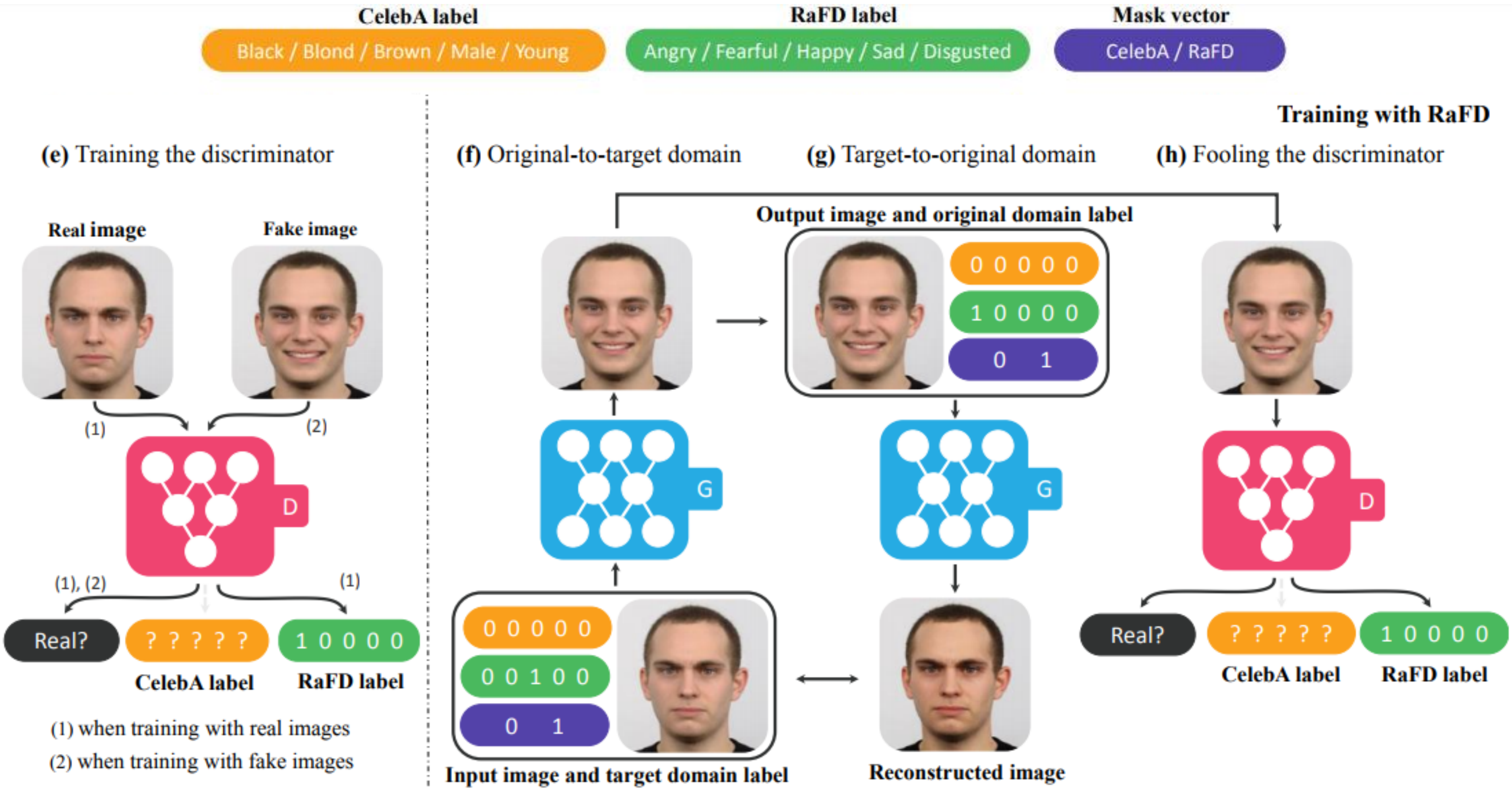
(a) Training the discriminator
(b) Original-to-target domain
(c) Target-to-original domain
(d) Fooling the discriminator
Real image
Fake image
(1)
(2)
D
(1), (2)
(1)
Real / Fake
Domain classification
Fake image
G
Target domain
Input image
Depth-wise concatenation
Depth-wise concatenation
Fake image
Original domain
G
Reconstructed image
Fake image
D
Real / Fake
Domain classification

# StarGAN



CelebA label
Black / Blond / Brown / Male / Young
RaFD label
Angry / Fearful / Happy / Sad / Disgusted
Mask vector
CelebA / RaFD
(a) Training the discriminator
(b) Original-to-target domain
(c) Target-to-original domain
(d) Fooling the discriminator
Real image
Fake image
(1)
(2)
D
(1), (2)
(1)
Real?
0 0 1 0 1
? ? ? ? ?
CelebA label
RaFD label
(1) when training with real images
(2) when training with fake images
Output image and original domain label
0 0 1 0 1
0 0 0 0 0
1 0
G
G
1 0 0 1 1
0 0 0 0 0
1 0
Input image and target domain label
Reconstructed image
D
Real?
1 0 0 1 1
? ? ? ? ?
CelebA label
RaFD label
Training with CelebA

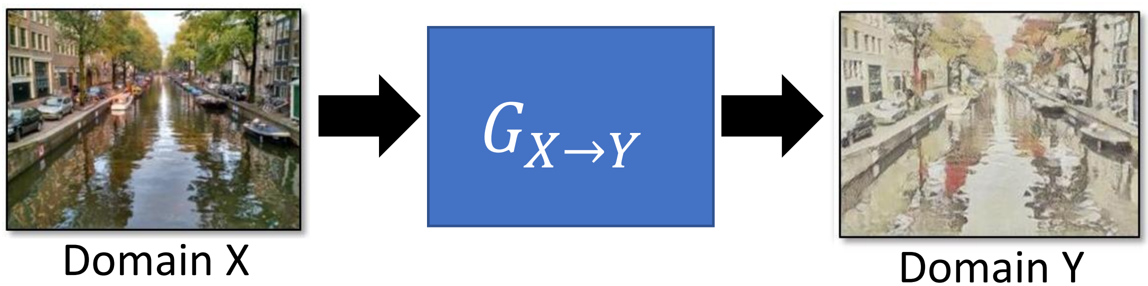
# StarGAN



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
Real image
Fake image
(1)
(2)
D
(1), (2)
(1)
Real?
? ? ? ? ?
1 0 0 0 0
CelebA label
RaFD label
(1) when training with real images
(2) when training with fake images
Output image and original domain label
0 0 0 0 0
1 0 0 0 0
0 1
G
G
D
0 0 0 0 0
0 0 1 0 0
0 1
Input image and target domain label
Reconstructed image
Real?
? ? ? ? ?
1 0 0 0 0
CelebA label
RaFD label

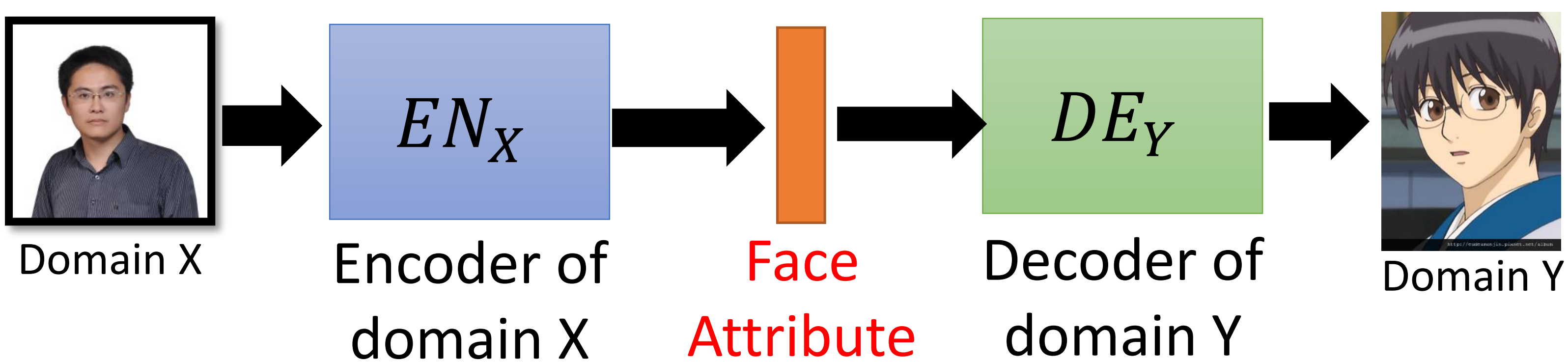
# Unsupervised Conditional Generation

- Approach 1: Direct Transformation




For texture or color change

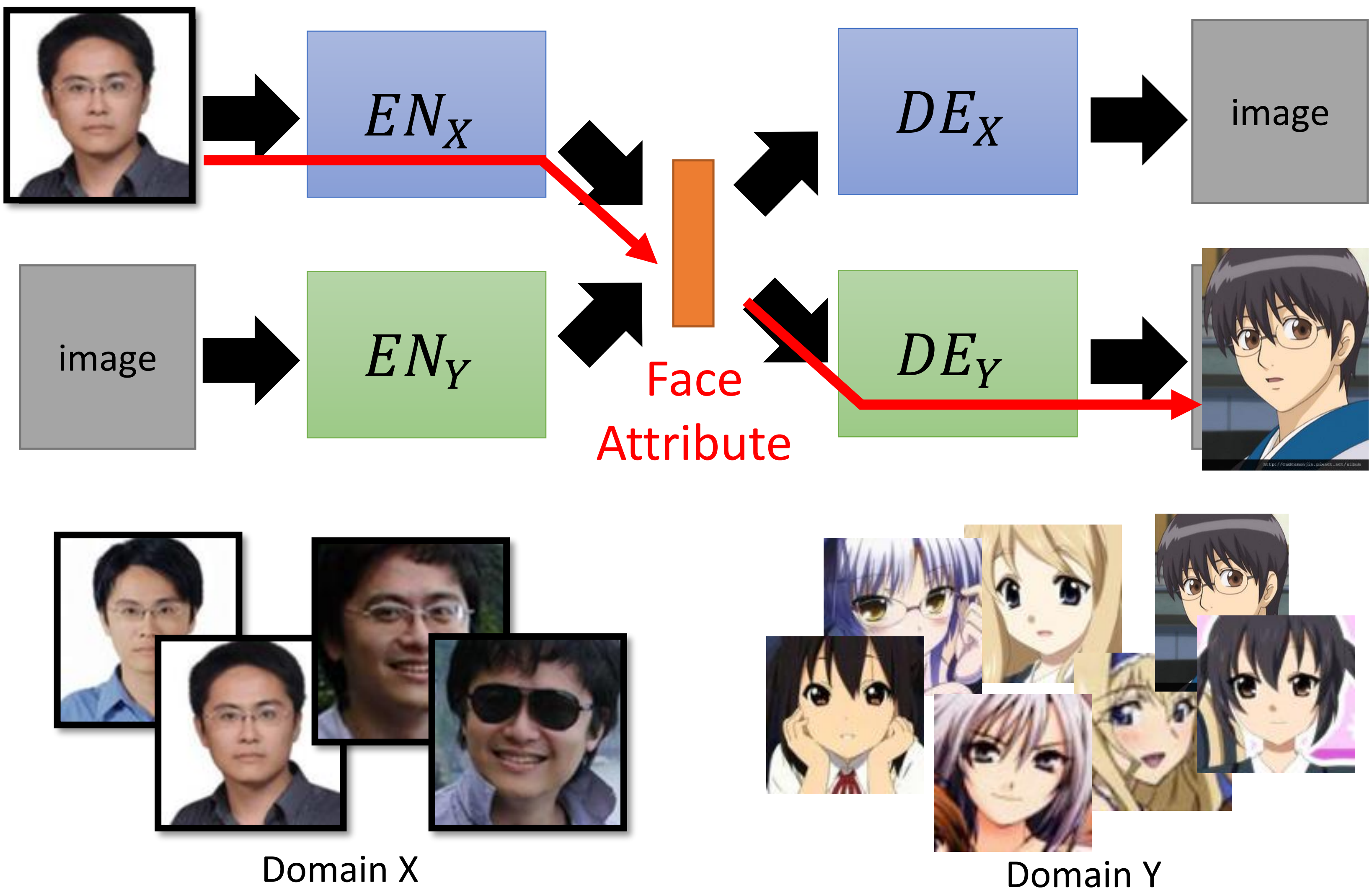
- Approach 2: Projection to Common Space




Larger change, only keep the semantics

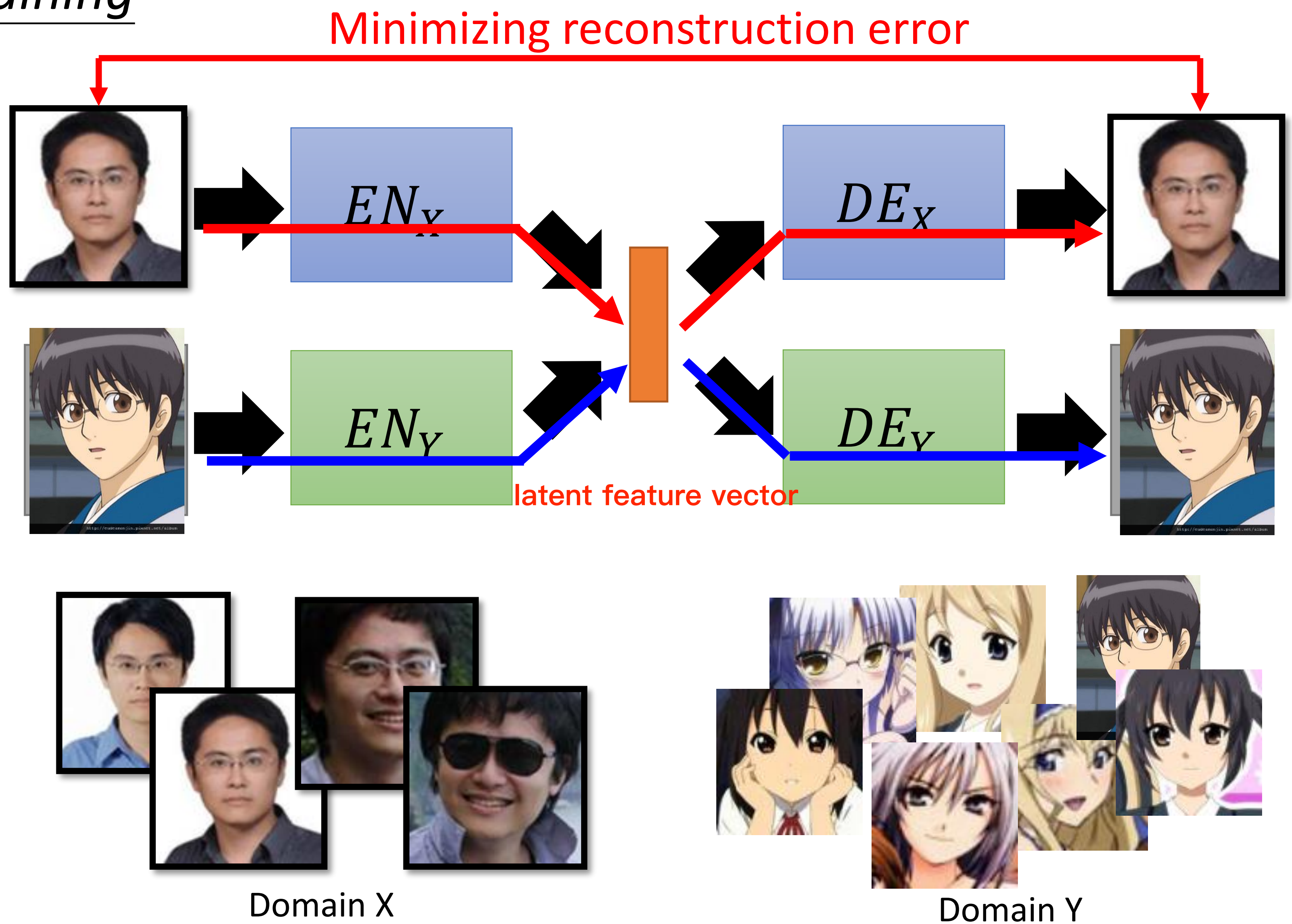
# Projection to Common Space

## Target



# *Projection to Common Space*

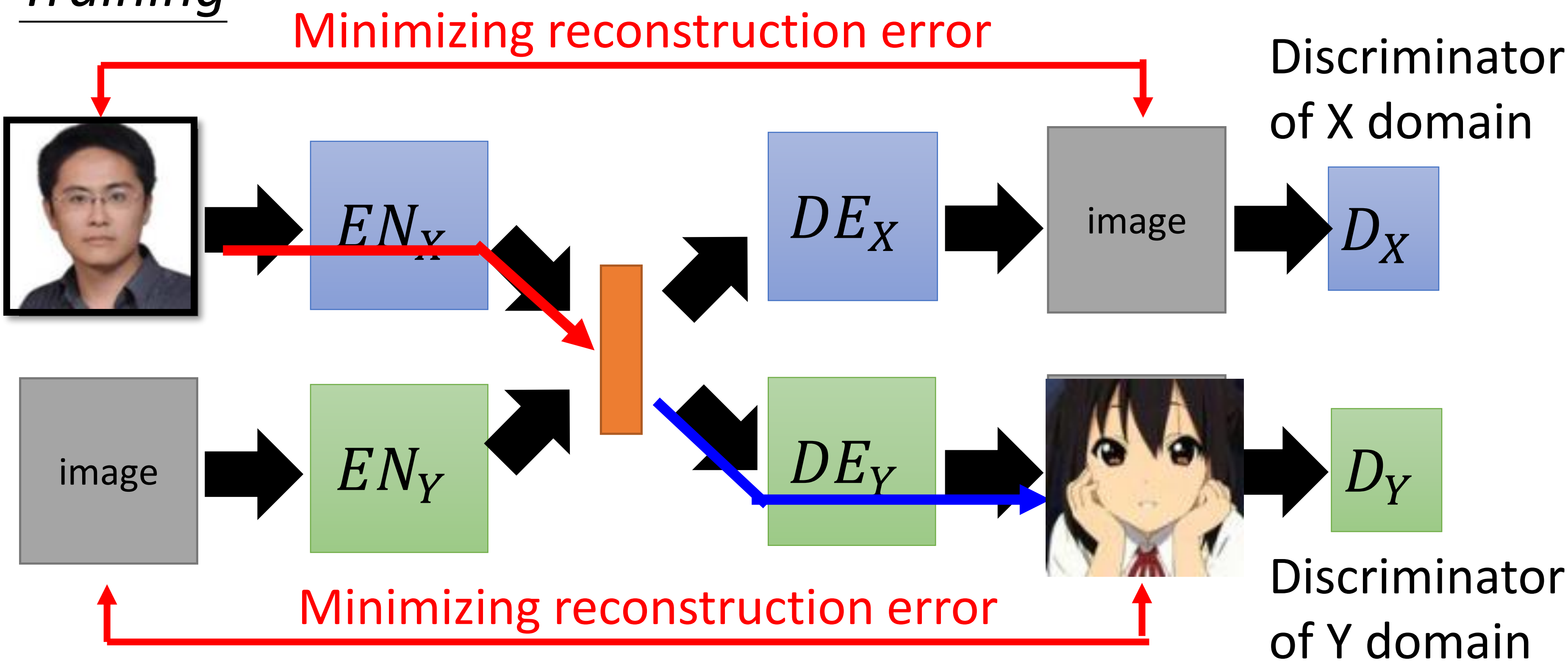
## *Training*



# *Projection to Common Space*

*Training*

加入discriminator判斷realistic

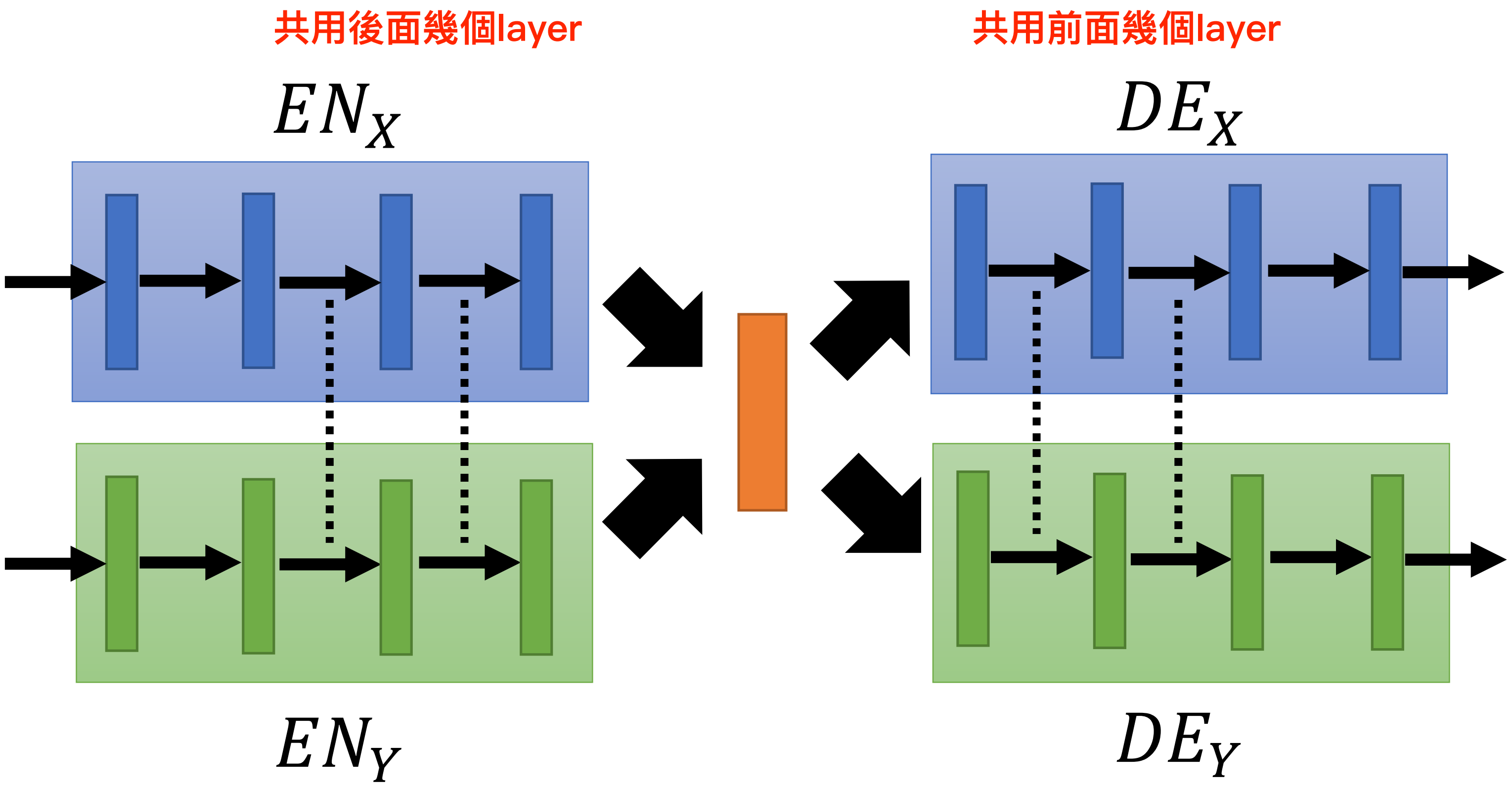



Because we train two auto-encoders separately …

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

# *Projection to Common Space*

## *Training*




甚至全部都共用參數，只是多位進去一個flag判斷domain

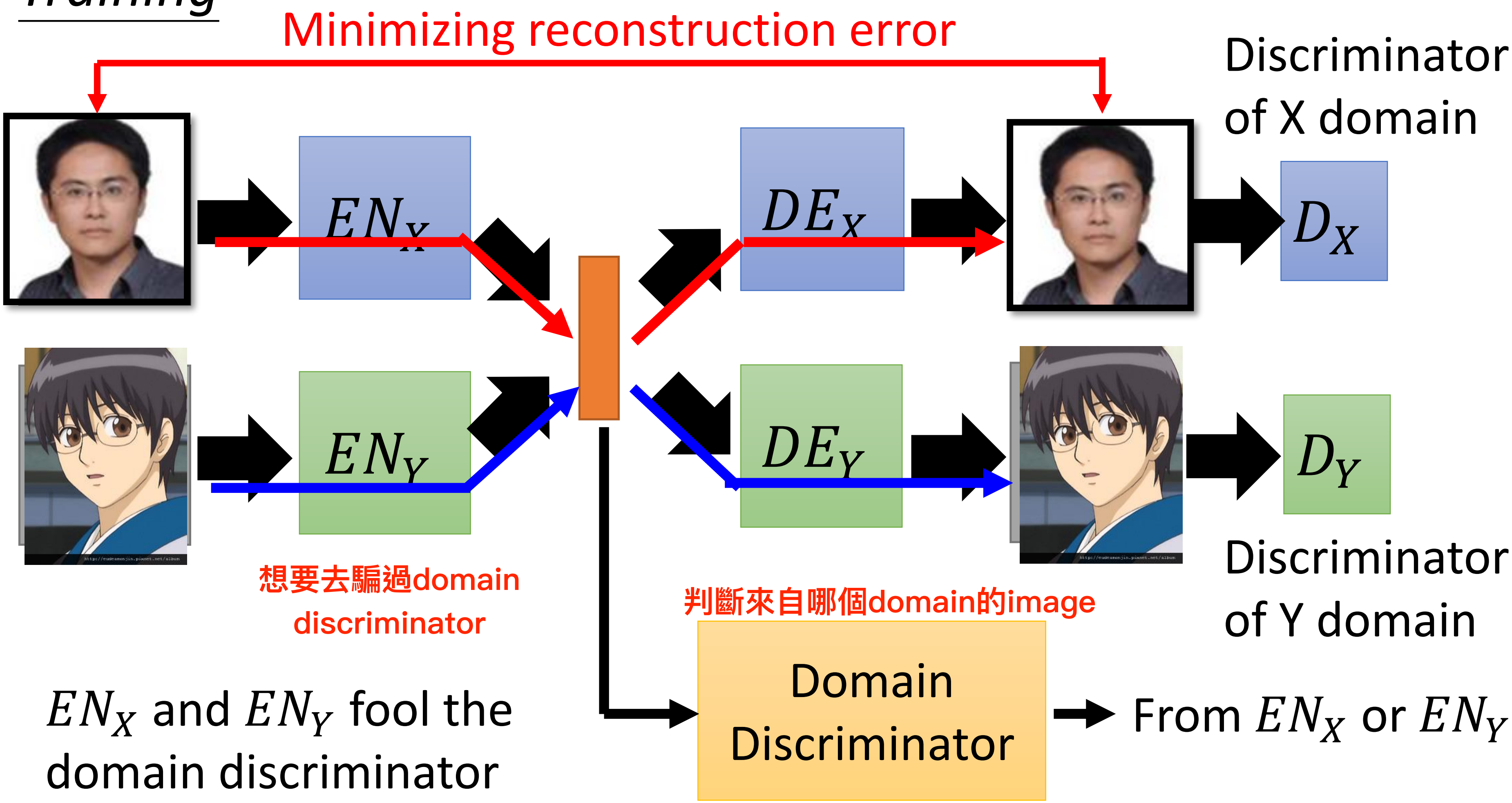
Sharing the parameters of encoders and decoders

Couple GAN[Ming-Yu Liu, et al., NIPS, 2016]

UNIT[Ming-Yu Liu, et al., NIPS, 2017]

# ***Projection to Common Space***

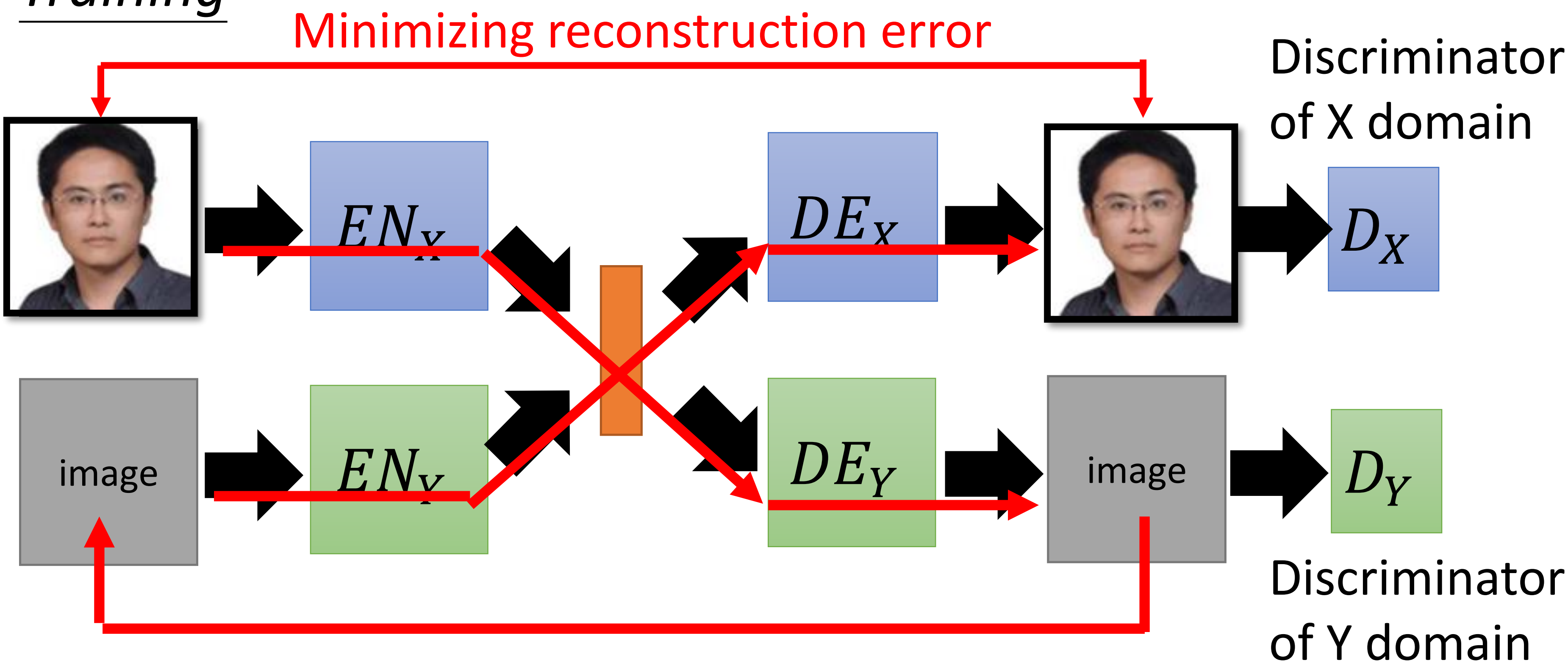
*Training*




The domain discriminator forces the output of $EN_X$ and $EN_Y$ have the same distribution. [Guillaume Lample, et al., NIPS, 2017]

# *Projection to Common Space*

## *Training*



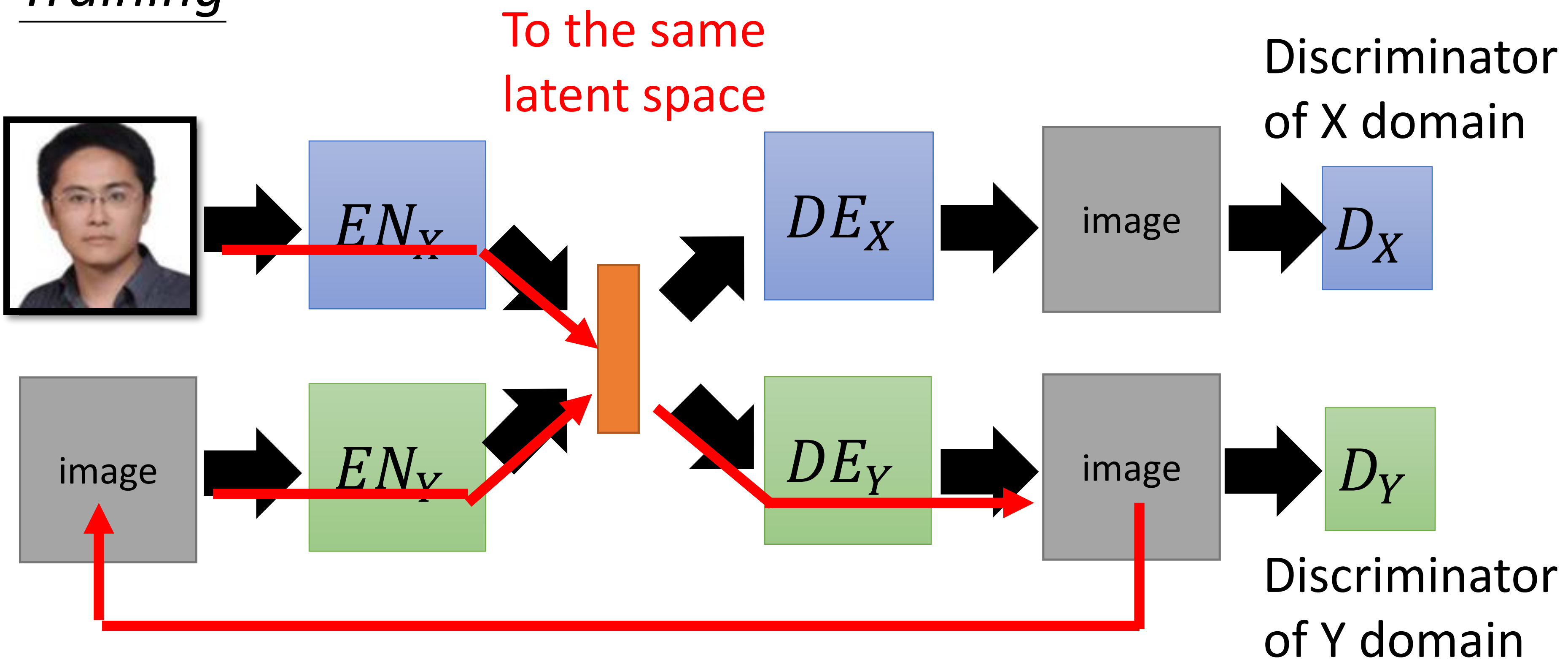

跟cycleGAN意思一樣的！只是這邊把Generator拆成encoder–decoder，並且對這些作組合

Cycle Consistency:

Used in ComboGAN [Asha Anoosheh, et al., arXiv, 017]

# *Projection to Common Space*

## *Training*




Semantic Consistency:

Used in DTN [Yaniv Taigman, et al., ICLR, 2017] and XGAN [Amélie Royer, et al., arXiv, 2017]

# UNIT: Unsupervised Image-to-image Translation

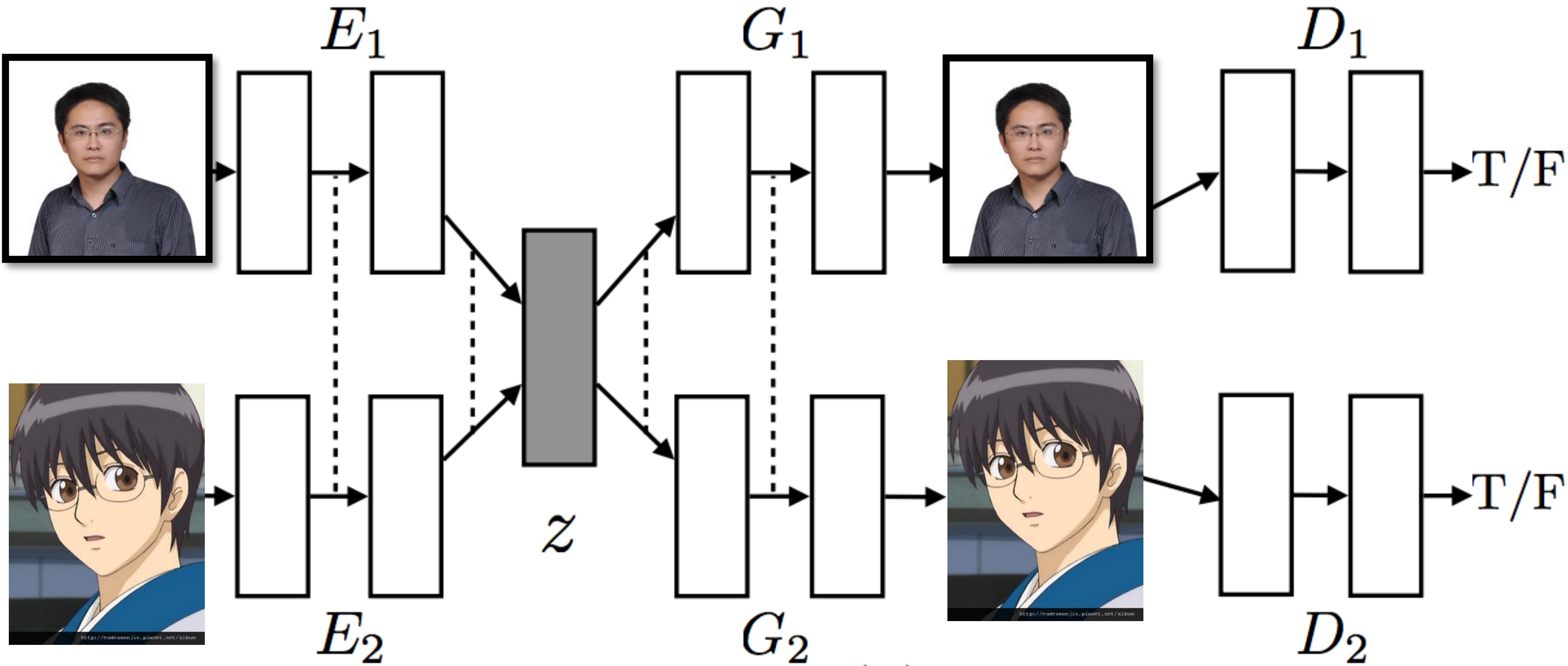



Table 1: Interpretation of the roles of the subnetworks in the proposed framework.

| Networks | $\{E_1, G_1\}$ | $\{E_1, G_2\}$ | $\{G_1, D_1\}$ | $\{E_1, G_1, D_1\}$ | $\{G_1, G_2, D_1, D_2\}$ |
|---|---|---|---|---|---|
| Roles | VAE for $\mathcal{X}_1$ | Image Translator $\mathcal{X}_1 \to \mathcal{X}_2$ | GAN for $\mathcal{X}_1$ | VAE-GAN [14] | CoGAN [17] |

https://arxiv.org/pdf/1703.00848.pdf
https://github.com/mingyuliutw/unit

- http://paintstransfer.com
- https://www.bilibili.com/video/av17537429/