By Bruce Lee

# StarGAN v2:
# Diverse Image Synthesis for Multiple Domains

*Optical Information System Laboratory*


李杰恩

**Chieh-En Lee**

celee@nctu.edu.tw

Department of Photonics

National Chiao Tung University

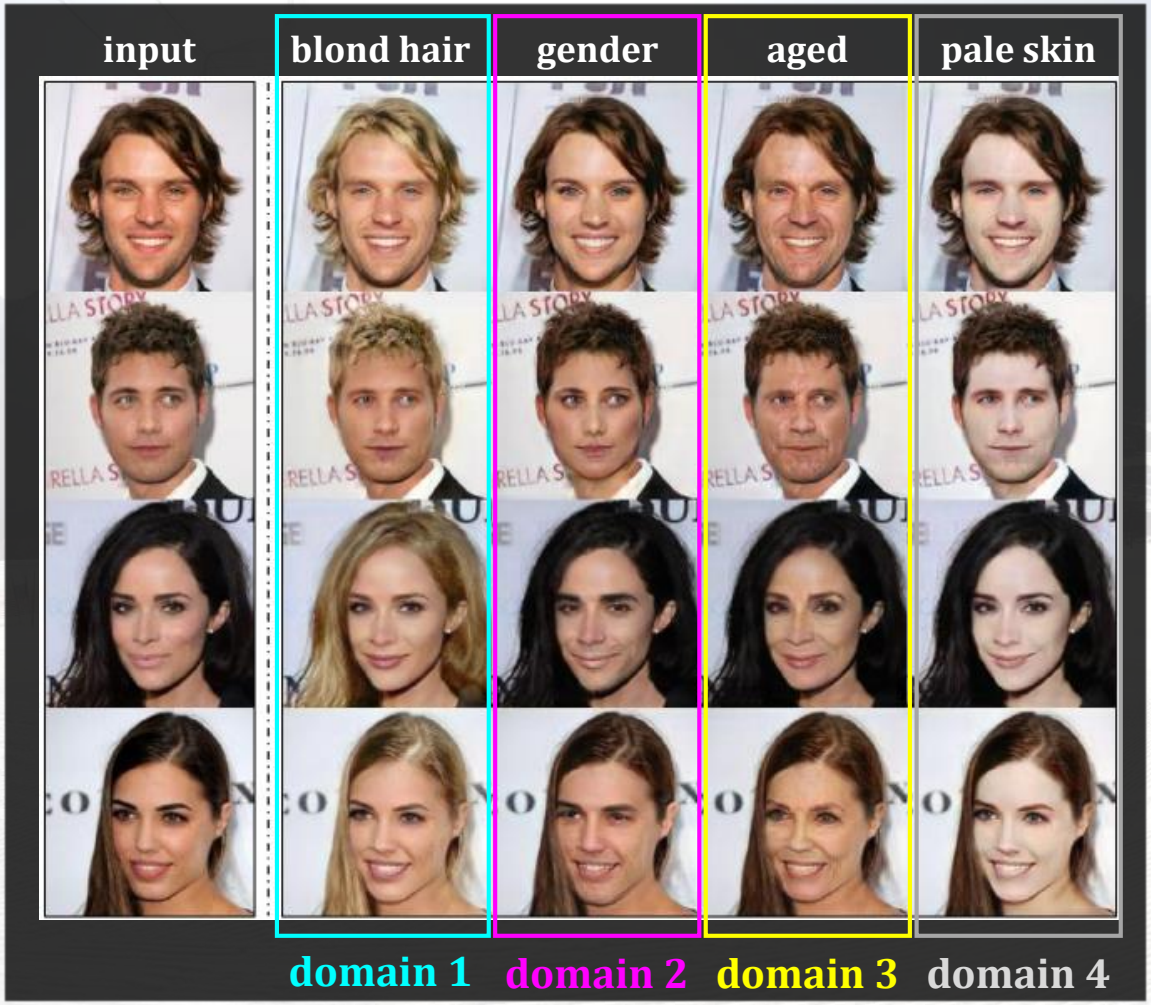
30010 Hsinchu, Taiwan

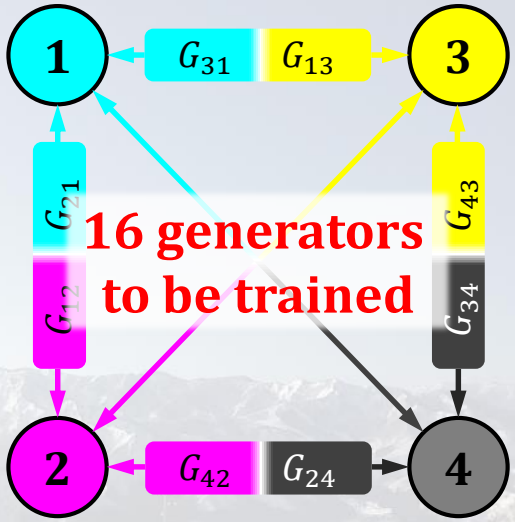

**Jan. 12th, 2021**

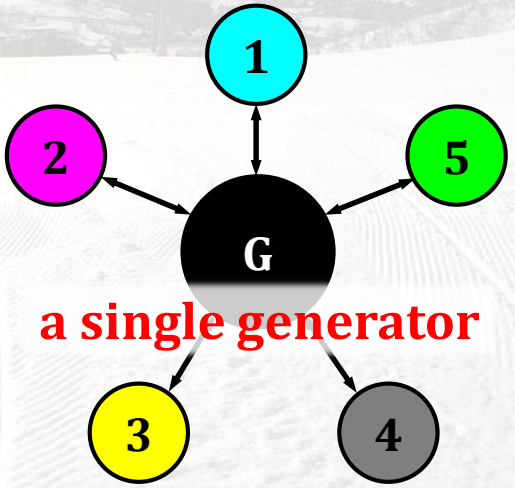
# Once upon a time…



- StarGAN v1 in 2018 IEEE/CVF CVPR



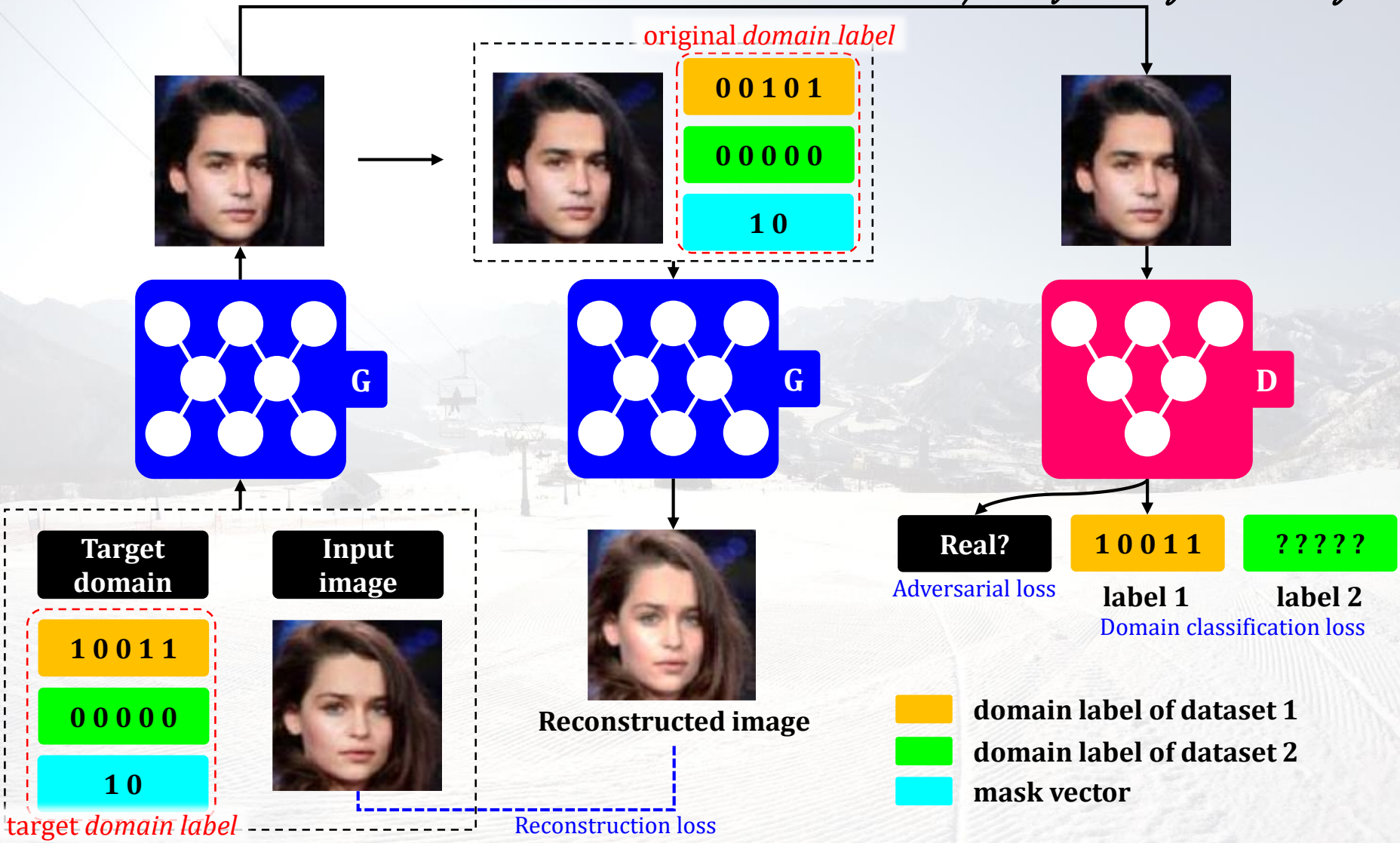




**Ref:** Y. Choi, *et al.*, Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. In *CVPR*, 2018.

# StarGAN v1: Target *domain label* and *mask vector*



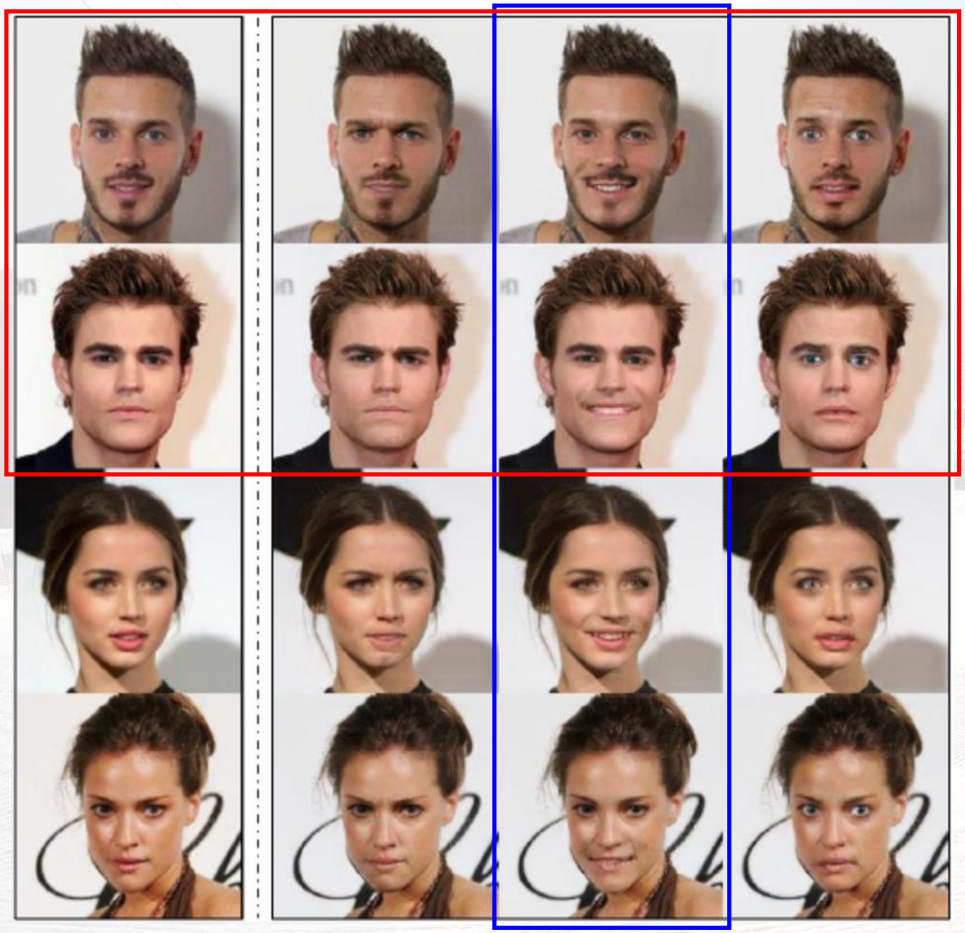



**Ref:** Y. Choi, *et al.*, Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. In *CVPR*, 2018.

# Here comes the problem



- Domain implies a set of images that can be grouped as visually distinctive category with *different styles*.



In StarGAN v1, this is a single domain *i*

In StarGAN v2, they should all be grouped as a single domain *i* (e.g., based on gender). Each image has a unique appearance called *style*. In this case, the *domain* is *Male*, and there are different styles *Angry*/*Happy*/*Fearful*.

StarGAN v1: **domain label**

StarGAN v2: **style code**

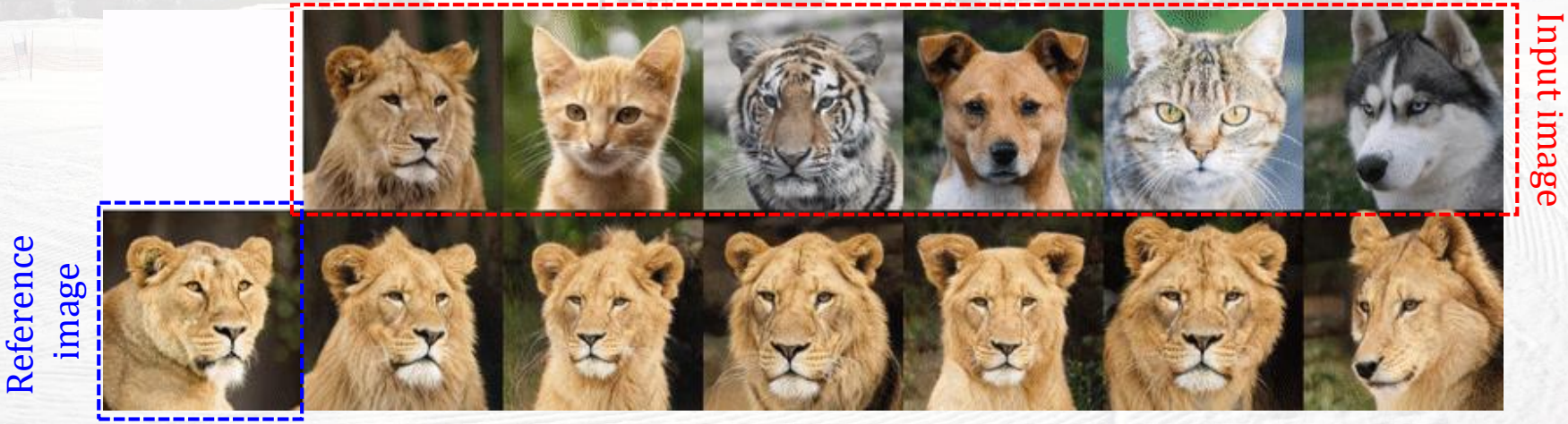
**Ref:** Y. Choi, *et al.*, Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. In *CVPR*, 2018.

# Before we start, there is one thing you SHOULD know



## 1. Latent-guided image synthesis

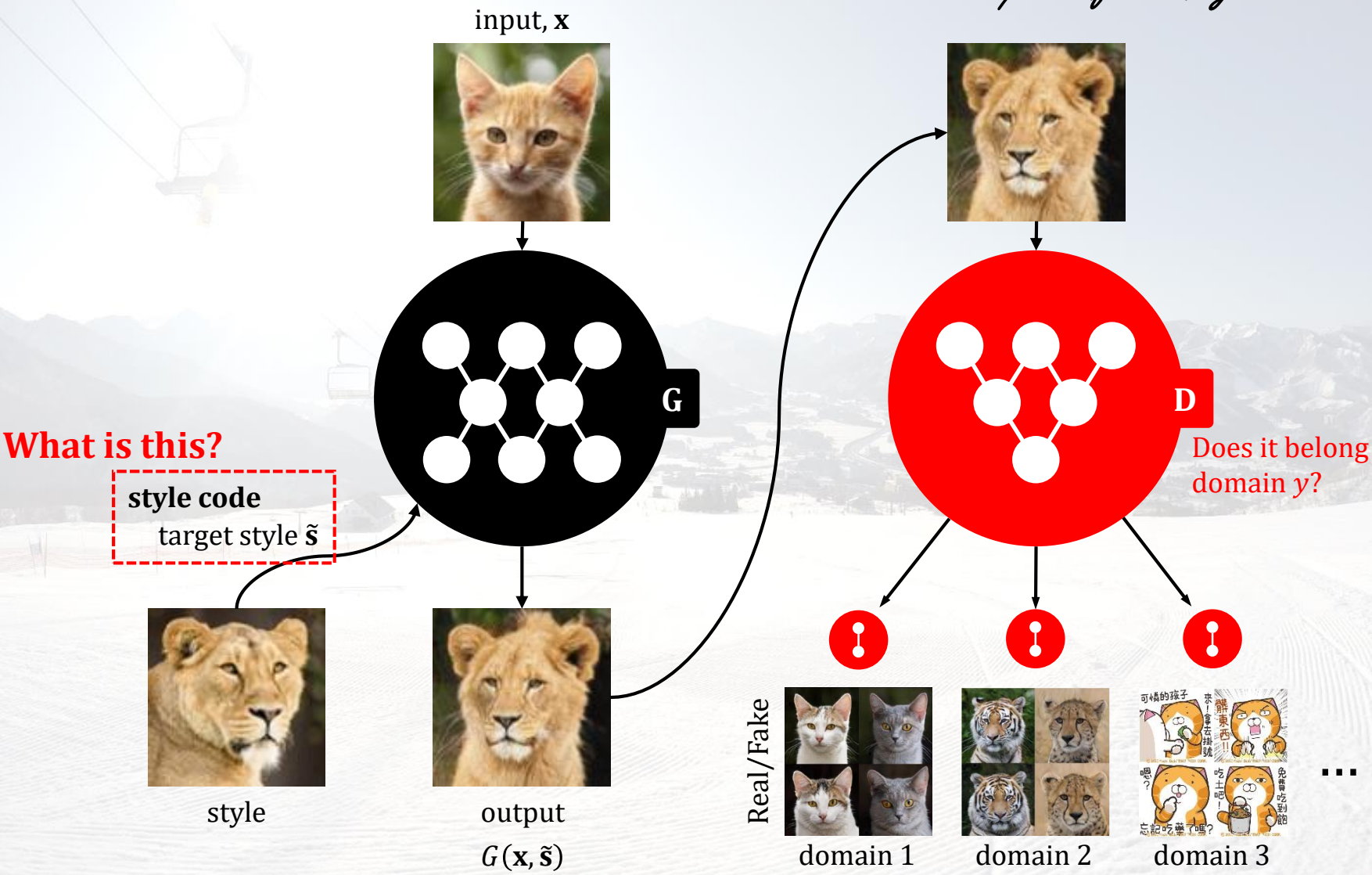

## 2. Reference-guided image synthesis




**Ref:** Y. Choi, *et al.*, StarGAN v2: Diverse image synthesis for multiple domains. In *CVPR*, 2020.

# Overview of StarGAN v2






**Ref:** Y. Choi, *et al.*, StarGAN v2: Diverse image synthesis for multiple domains. In *CVPR*, 2020.

# StarGAN v2: *Mapping network* and *style encoder*



## Mapping network

generated by random Gaussian noise

$\mathbf{z}_i$ **Latent code**

$i = 1, 2, 3, \ldots$

$\mathbf{s}_{iy} = F_y(\mathbf{z}_i)$

**F**

style code in domain 1 $\mathbf{s}_{i1}$

style code in domain 2 $\mathbf{s}_{i2}$

style code in domain 3 $\mathbf{s}_{i3}$

## Style encoder

reference image $\mathbf{x}_{ref}$ with known original domain

$\mathbf{s}_y = E_y(\mathbf{x}_{ref})$

**E**

style code in domain 1 $\mathbf{s}_1$

style code in domain 2 $\mathbf{s}_2$

style code in domain 3 $\mathbf{s}_3$

...

**Ref:** Y. Choi, *et al.*, StarGAN v2: Diverse image synthesis for multiple domains. In *CVPR*, 2020.

# Concept of latent-guided image synthesis



## Mapping network

generated by random Gaussian noise

$\mathbf{z}_i$ **Latent code**

$i = 1, 2, 3, \ldots$

$\mathbf{s}_{iy} = F_y(\mathbf{z}_i)$

**F**

style code in domain 1 $\mathbf{s}_{i1}$

style code in domain 2 $\mathbf{s}_{i2}$

style code in domain $y$ $\mathbf{s}_{iy}$

randomly selected a style code of domain $y$ as the target style $\tilde{\mathbf{s}}$

**style code**

target style $\tilde{\mathbf{s}}$

no reference in this case

input, **x**

**G**


output $G(\mathbf{x}, \tilde{\mathbf{s}})$

domain $y$

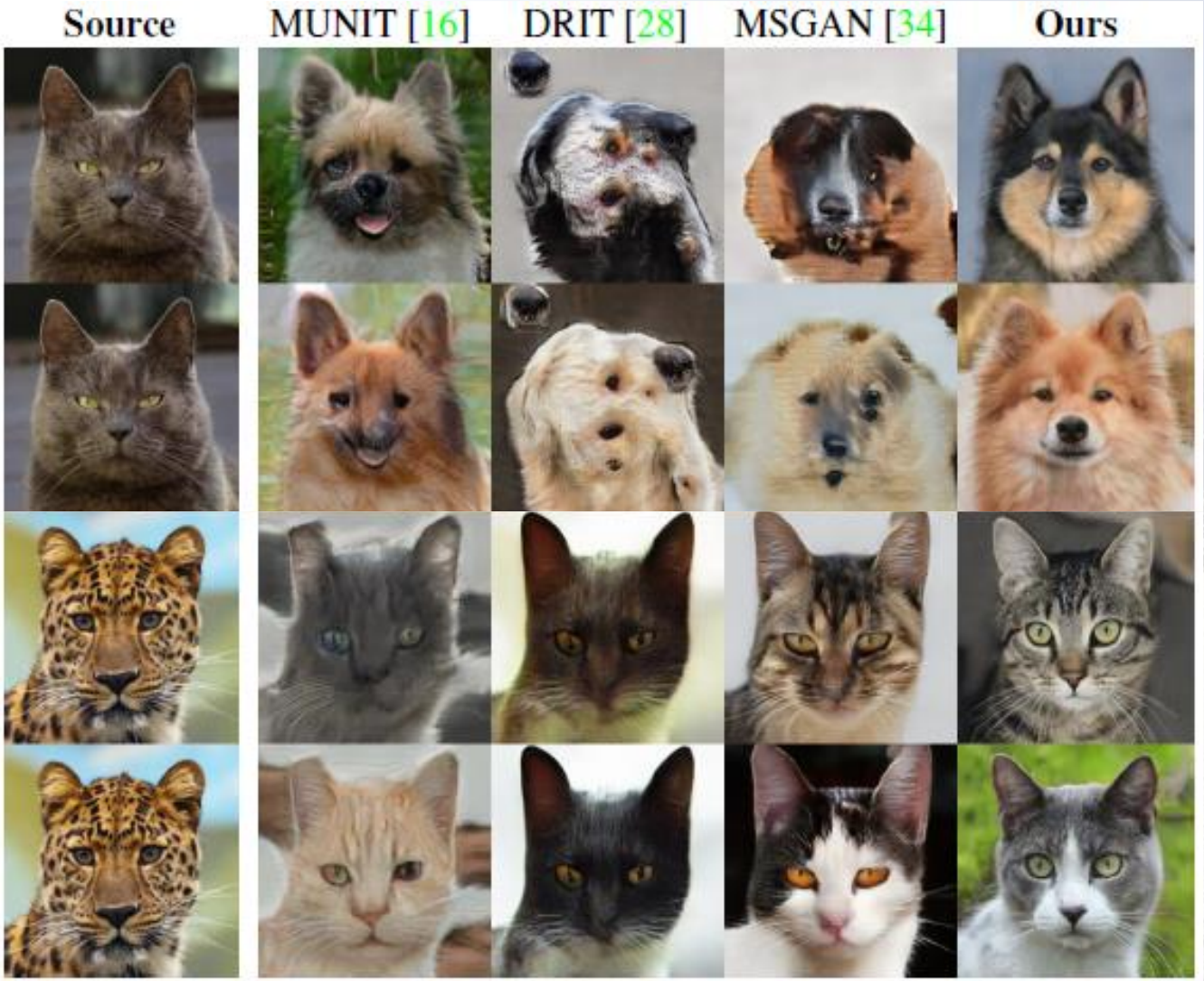

**D**

Real or Fake?

**Ref:** Y. Choi, *et al.*, StarGAN v2: Diverse image synthesis for multiple domains. In *CVPR*, 2020.

# Latent-guided image synthesis results






**Ref:** Y. Choi, *et al.*, StarGAN v2: Diverse image synthesis for multiple domains. In *CVPR*, 2020.

# Concept of reference-guided image synthesis



## Style encoder

reference image $\mathbf{x}_{ref}$ with known original domain

$$\mathbf{s}_y = E_y(\mathbf{x}_{ref})$$

E

encoding the reference image into a style code of domain $y$ as the target style $\tilde{\mathbf{s}}$

style code in domain 1 $\mathbf{s}_1$

style code in domain 2 $\mathbf{s}_2$

style code in domain 3 $\mathbf{s}_3$

...

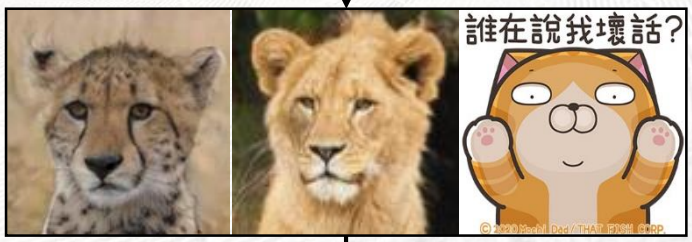
**style code**

target style $\tilde{\mathbf{s}}$

input, $\mathbf{x}$

G

output $G(\mathbf{x}, \tilde{\mathbf{s}})$




$G(\mathbf{x}, \mathbf{s}_1)$ $G(\mathbf{x}, \mathbf{s}_2)$ $G(\mathbf{x}, \mathbf{s}_3)$

D Real or Fake?

**Ref:** Y. Choi, *et al.*, StarGAN v2: Diverse image synthesis for multiple domains. In *CVPR*, 2020.

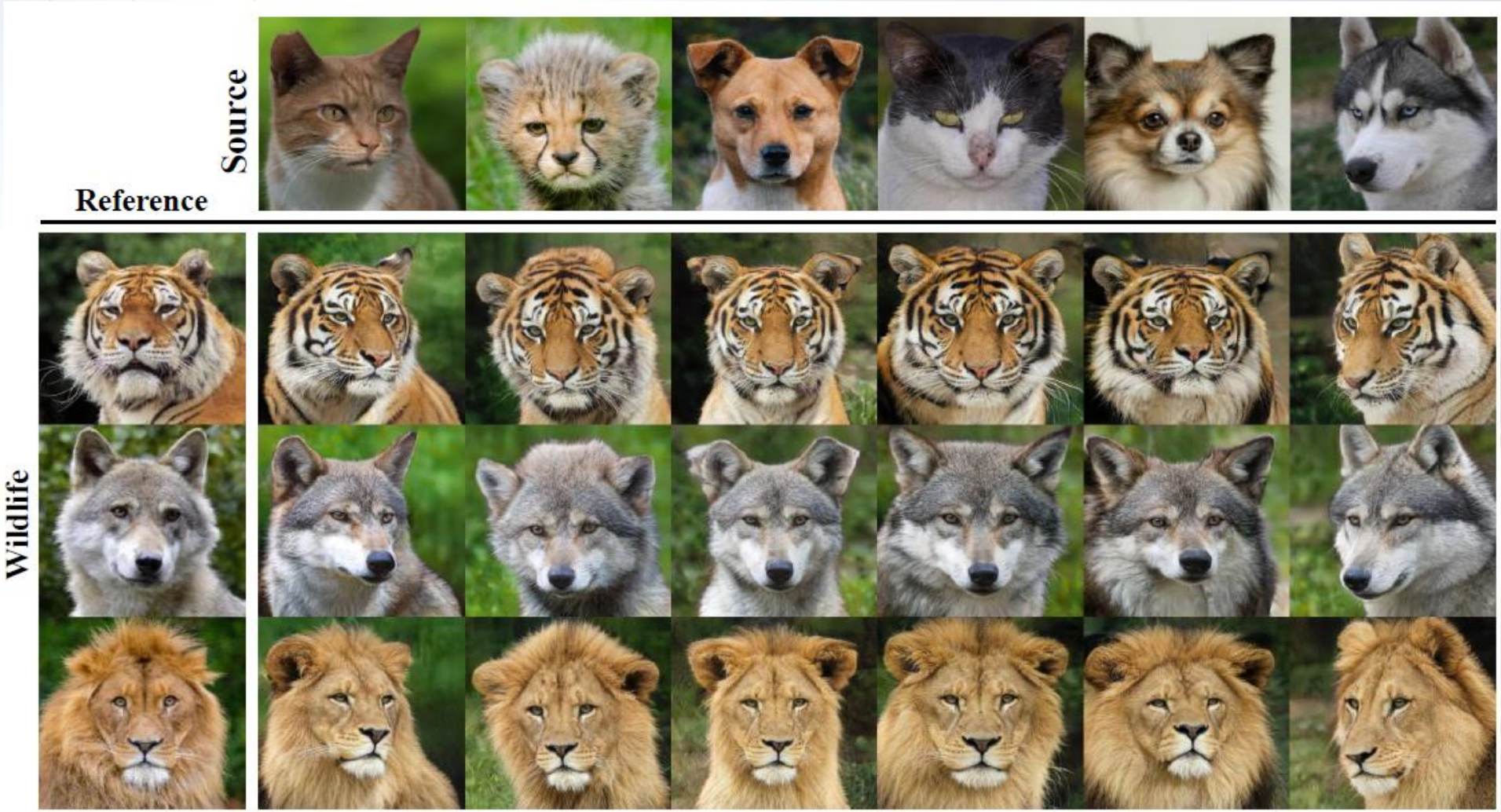
# Reference-guided image synthesis results (1/2)




**Ref:** Y. Choi, *et al.*, StarGAN v2: Diverse image synthesis for multiple domains. In *CVPR*, 2020.

# Reference-guided image synthesis results (2/2)



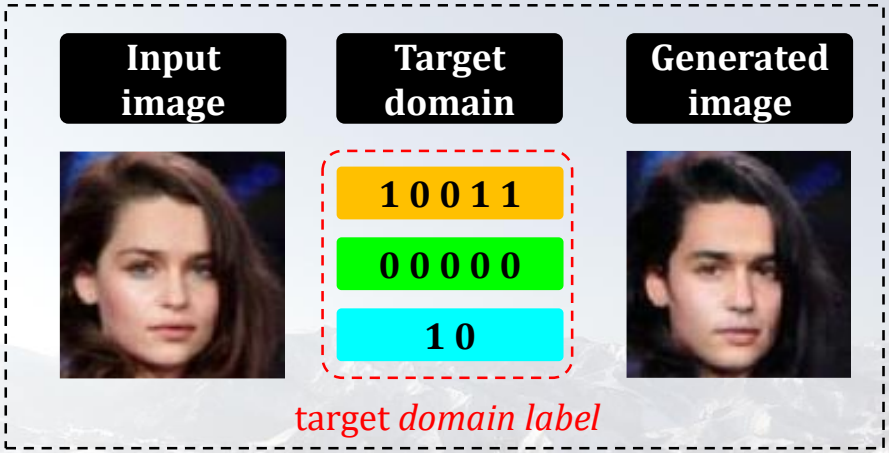



**Ref:** Y. Choi, *et al.*, StarGAN v2: Diverse image synthesis for multiple domains. In *CVPR*, 2020.

# Summary



- StarGAN v1:
  - Multi-domain image translation
  - Single generator
  - The idea of domain label
  - Mask vector
  - Lack of domain diversity




- StarGAN v2:
  - Multi-domain image translation
  - Single generator
  - Diversity of each domain
  - The idea of style code
  - Mapping network and style encoder


***Thank you!***

***Please feel free to contact me if you have any question or comment.***

***Jie-En Li:***

*celee@nctu.edu.tw*