

StarGAN: Unified generative adversarial networks for multi-domain image-to- image translation

2017, Yunjey Choi et al.

한양대학교
컴퓨터소프트웨어학과
인공지능 연구실
조건희

**Unified generative
adversarial networks for
multi-domain image-to-
image translation**

**Unified generative
adversarial networks for
multi-domain image-to-
image translation**

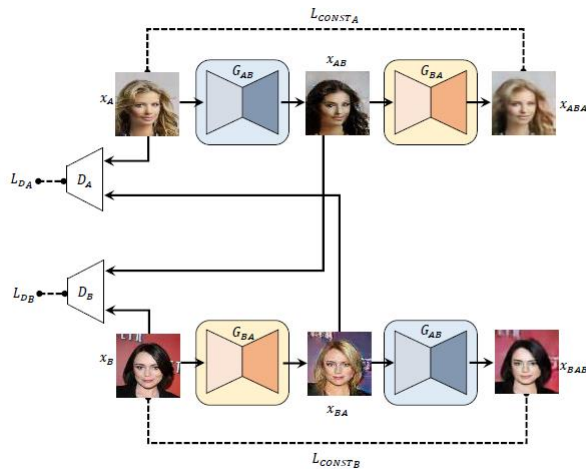
Unified generative
adversarial networks for
**multi-domain image-to-
image translation**

Unified generative
adversarial networks for
multi-domain image-to-
image translation

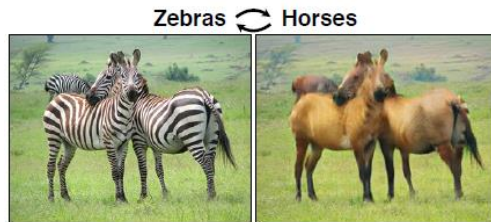


때용?

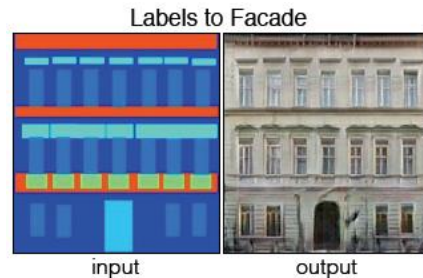
- 이미지 변환과 관련된 기존의 연구들 (DiscoGAN, CycleGAN, pix2pix 등)은 도메인 간 이미지 변환에서 좋은 성능을 보여줌.



DiscoGAN



CycleGAN



pix2pix

※ “Domain” 은 같은 attribute value를 공유하는 이미지 집합

- 그러나 기존의 모델들은 모두 2개의 도메인 간 변환에만 집중하였음.
- 따라서, Multi-domain으로 문제를 확장하면 기존 모델들은 사용하기 어려움.

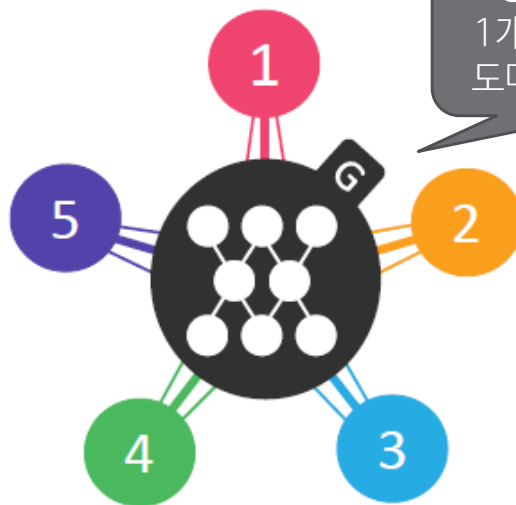
(a) Cross-domain models



기존 모델들은 여러 도메인 간 변환을 위해서는 각 도메인 pair 마다 G를 새로 학습해줘야 한다!

k개의 도메인이 있다면 만들어야 하는 G는 $k(k-1)$ 개

(b) StarGAN



StarGAN은 G를 1개만 학습하여 여러 도메인 간 변환 가능!

- 여러 도메인 간 이미지 변환의 예시

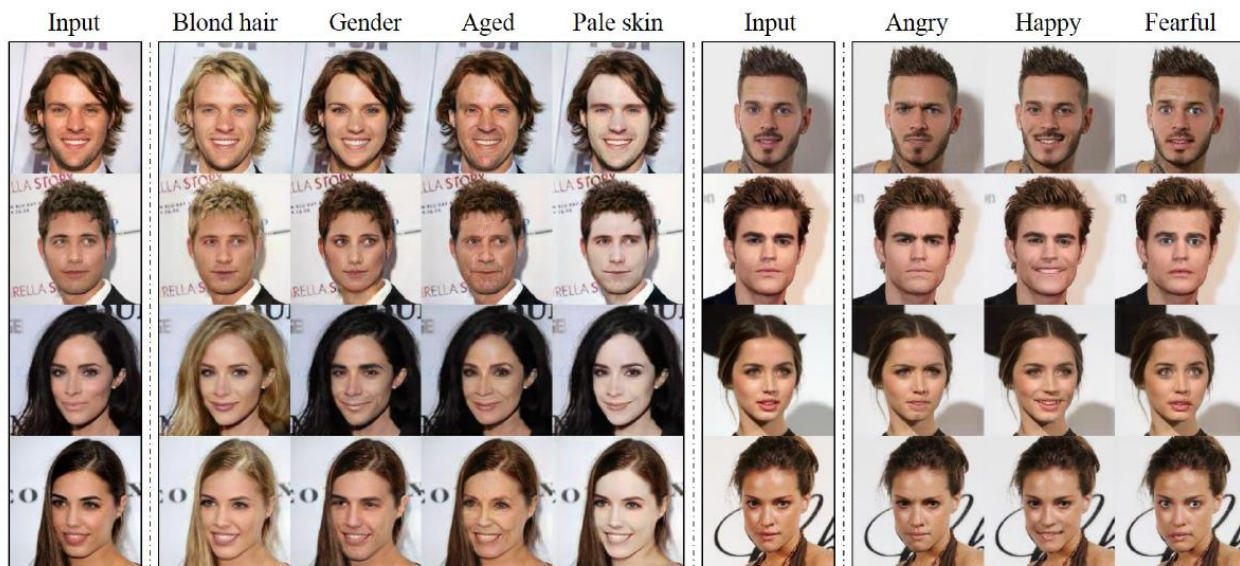
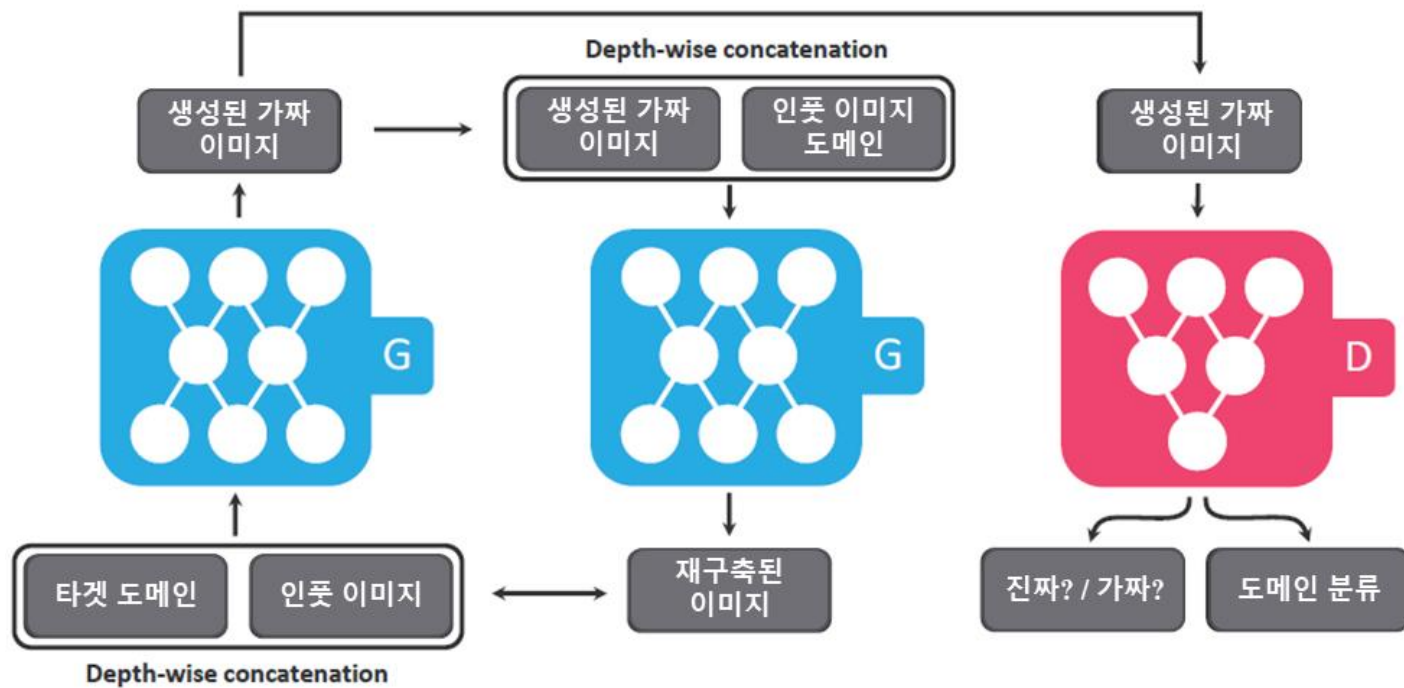


Figure 1. Multi-domain image-to-image translation results on the CelebA dataset via transferring knowledge learned from the RaFD dataset. The first and sixth columns show input images while the remaining columns are images generated by StarGAN. Note that the images are generated by a single generator network, and facial expression labels such as angry, happy, and fearful are from RaFD, not CelebA.

■ 전체 모델 구조



Discriminator 학습

CelebA Dataset



도메인 라벨

black hair/ blonde hair/ Male / Young

CelebA dataset



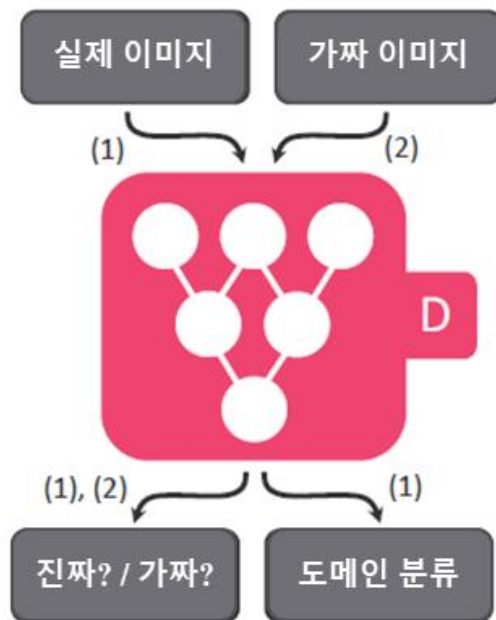
도메인 라벨

black hair/ blonde hair/ Male / Young

1 0 0 1

도메인 라벨

black hair/ blonde hair/ Male / Young



도메인 라벨

black hair/ blonde hair/ Male / Young

실제 이미지



(1)



(1), (2)

(1) Domain classification

진짜

1 0 0 1

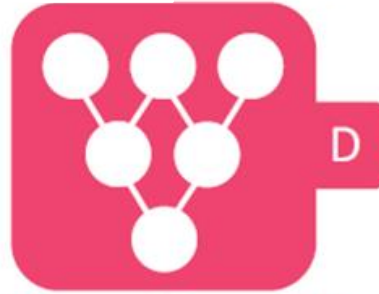
도메인 라벨

black hair/ blonde hair/ Male / Young

가짜 이미지



(2)



(1), (2)

가짜

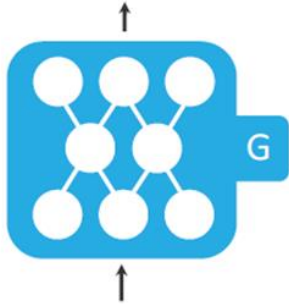
Generator 학습

도메인 라벨

black hair/ blonde hair/ Male / Young



생성된 이미지



0101

타겟 도메인



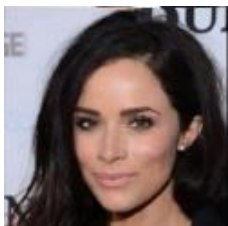
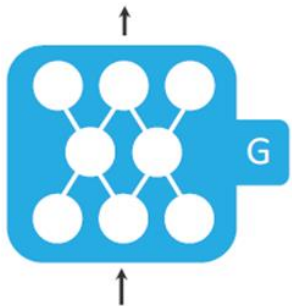
인풋 이미지

도메인 라벨

black hair/ blonde hair/ Male / Young



생성된 이미지



인풋 이미지

0101

타겟 도메인



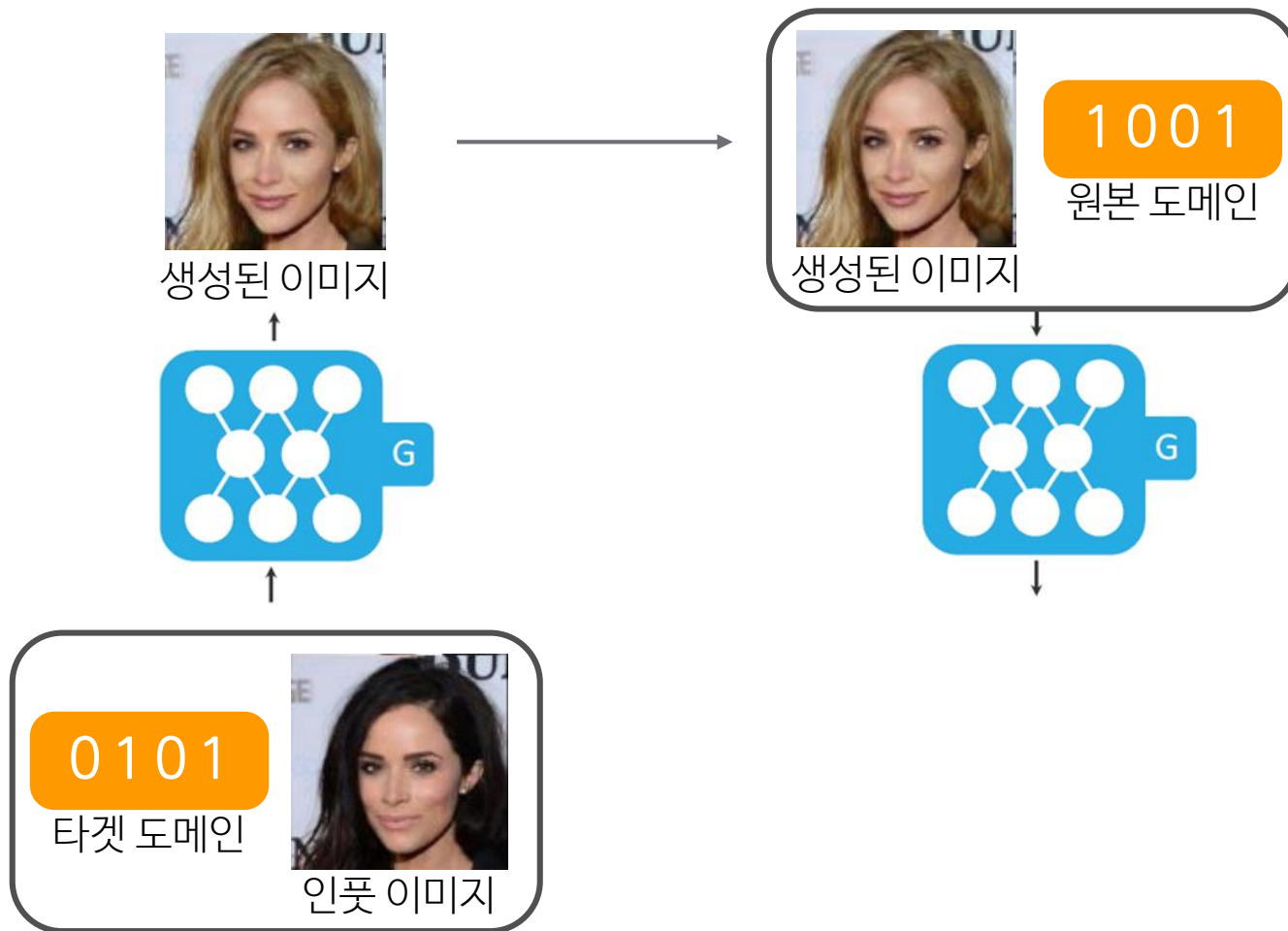
생성된 이미지

1001

원본 도메인

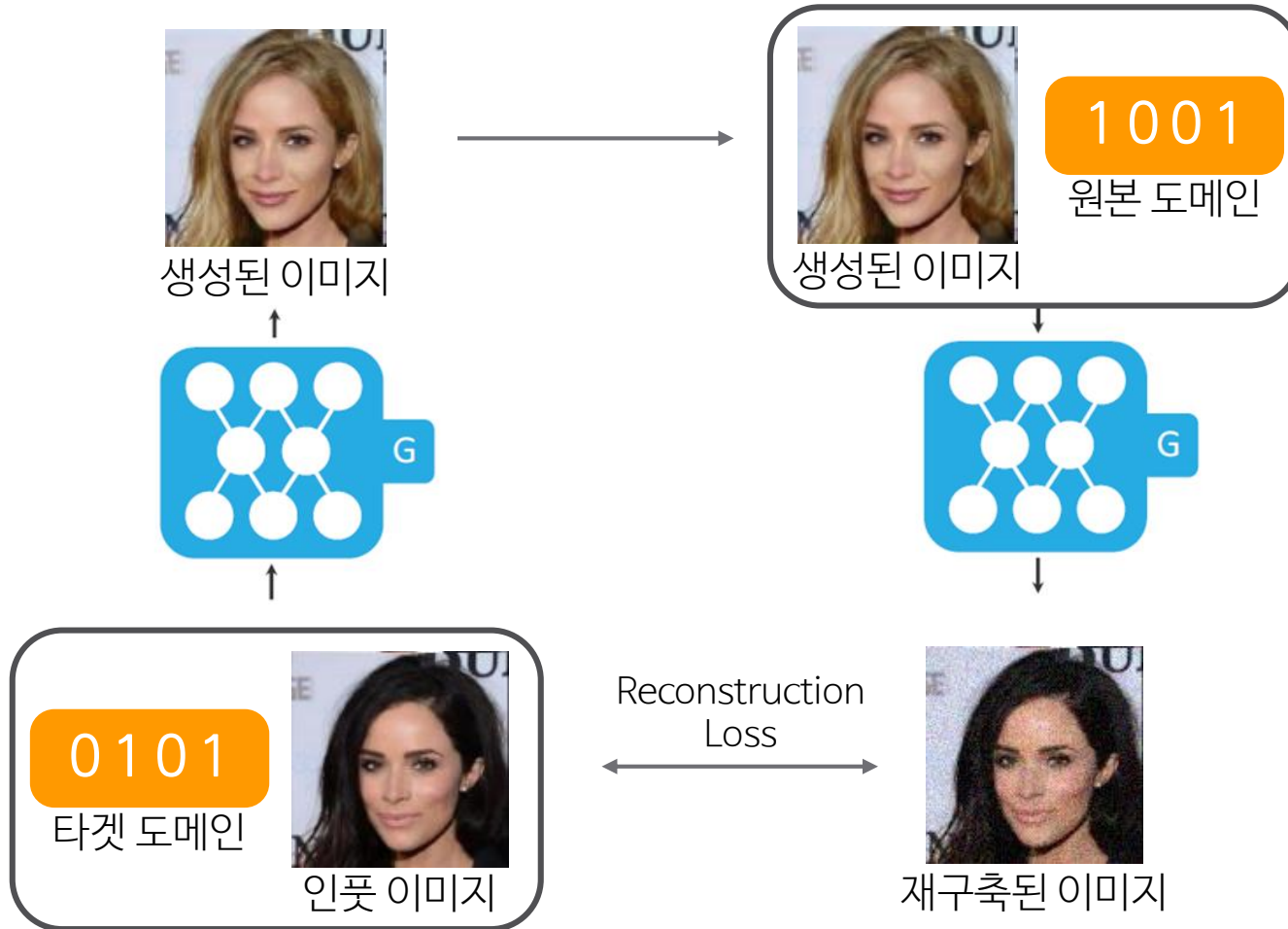
도메인 라벨

black hair/ blonde hair/ Male / Young



도메인 라벨

black hair/ blonde hair/ Male / Young



Reconstruction Loss가 왜 필요할까?

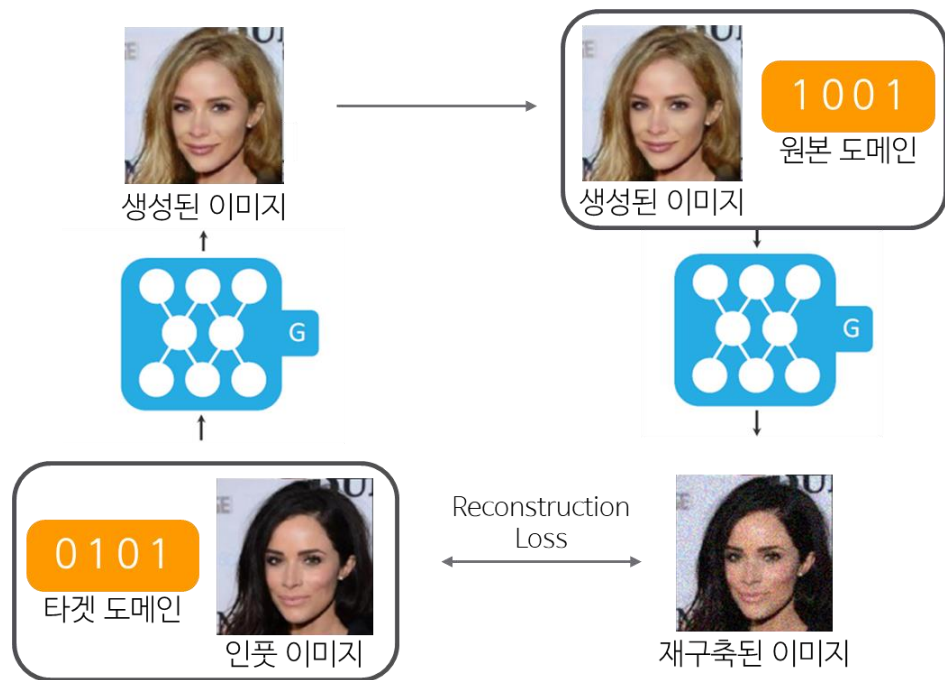
Reconstruction Loss가 왜 필요할까?

G가 생성한 이미지가 타겟 도메인으로
변환되면서도 **원본의 내용은 유지할 수**
있도록 하기 위해서!

다시 Discriminator 학습

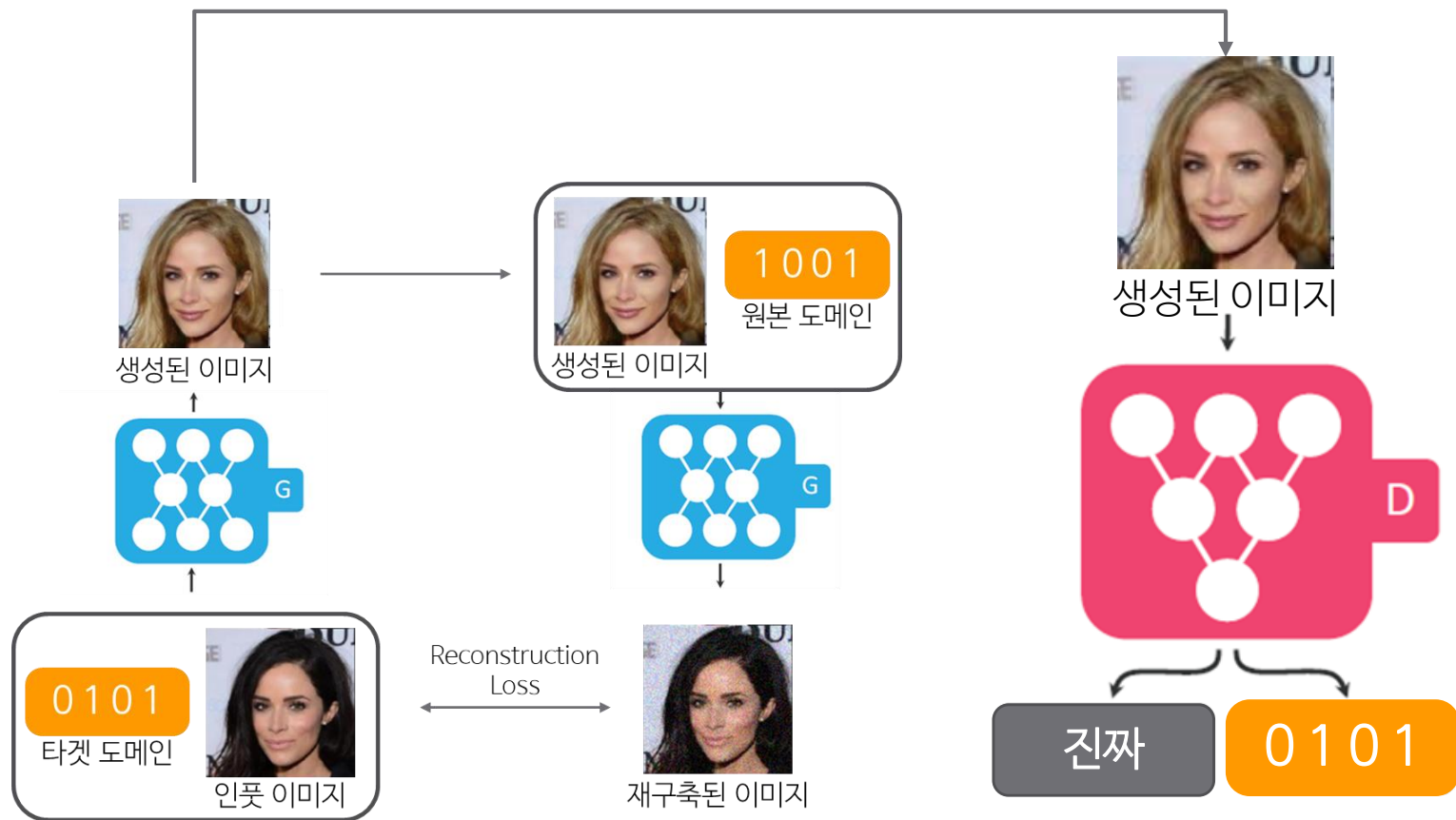
도메인 라벨

black hair/ blonde hair/ Male / Young



도메인 라벨

black hair/ blonde hair/ Male / Young



- Loss function
 - Adversarial Loss
 - Domain Classification Loss
 - Reconstruction Loss

- Adversarial Loss

- GAN의 기본적인 Loss 함수
- D 입장에서는 최대화하는 방향으로 학습
- G 입장에서는 최소화하는 방향으로 학습

$$\begin{aligned}\mathcal{L}_{adv} &= \mathbb{E}_x[\log D_{src}(x)] \\ &+ \mathbb{E}_{x,c}[\log(1 - D_{src}(G(x, c)))]\end{aligned}$$

- Domain Classification Loss
 - D에 추가적으로 붙어있는 Domain classifier를 학습하기 위해 사용되는 Loss 함수

$$\mathcal{L}_{cls}^r = \mathbb{E}_{x, c'} [-\log D_{cls}(c' | x)]$$

실제 이미지를 D에 넣었을 때 도메인 분류 결과와 실제 도메인과의 차이를 줄이는 방향으로 학습

$$\mathcal{L}_{cls}^f = \mathbb{E}_{x, c} [-\log D_{cls}(c | G(x, c))]$$

생성한 이미지를 D에 넣었을 때 도메인 분류 결과와 이미지를 생성할 때 인풋으로 넣었던 타겟 도메인과의 차이를 줄이는 방향으로 학습

- Reconstruction Loss
 - 원본 이미지의 domain-related part를 제외한 내용을 유지하며 학습하기 위한 Loss 함수

$$\mathcal{L}_{rec} = \mathbb{E}_{x,c,c'} [\|x - G(G(x, c), c')\|_1]$$

생성한 이미지와 인풋 이미지의 실제 도메인을 G에 인풋으로 넣었을 때 생성하는 이미지와 처음 인풋 이미지와의 차이(L1 norm)를 줄이는 방향으로 학습

- Full Objective

$$\mathcal{L}_D = -\mathcal{L}_{adv} + \lambda_{cls} \mathcal{L}_{cls}^r$$

$$\mathcal{L}_G = \mathcal{L}_{adv} + \lambda_{cls} \mathcal{L}_{cls}^f + \lambda_{rec} \mathcal{L}_{rec}$$

StarGAN의 놀라운 장점!

StarGAN의 놀라운 장점!

여러 가지 데이터셋을 동시에 학습 가능!

CelebA 데이터셋과 RaFD 데이터셋의 예를 들면,
이 2개의 데이터셋은 서로 도메인 라벨 정보가 다름

CelebA 데이터셋과 RaFD 데이터셋의 예를 들면,
이 2개의 데이터셋은 서로 도메인 라벨 정보가 다름

CelebA의 라벨 : ‘머리색’, ‘성별’, ‘나이’ 등

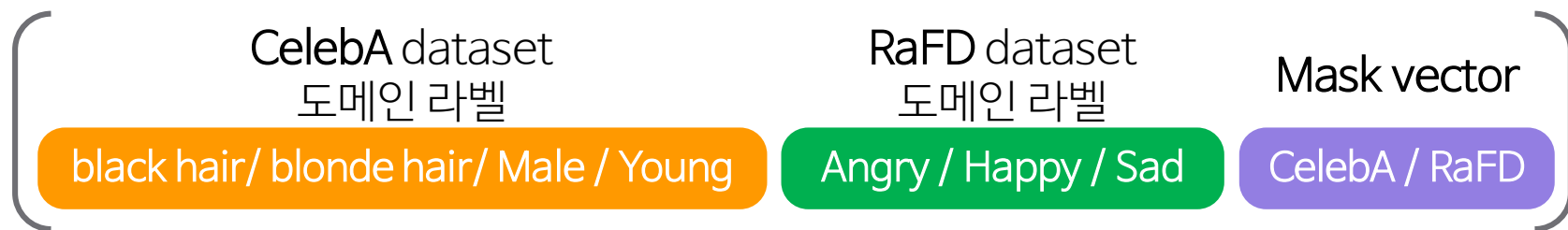
RaFD의 라벨 : ‘기쁨’, ‘슬픔’ 등 (표정)

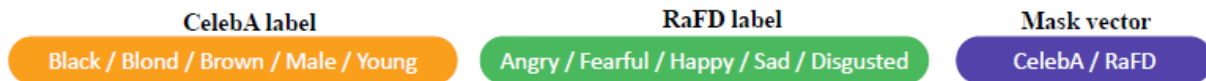
데이터셋마다 도메인 라벨 벡터 차원이 다르다!?

데이터셋마다 도메인 라벨 벡터 차원이 다르다!?

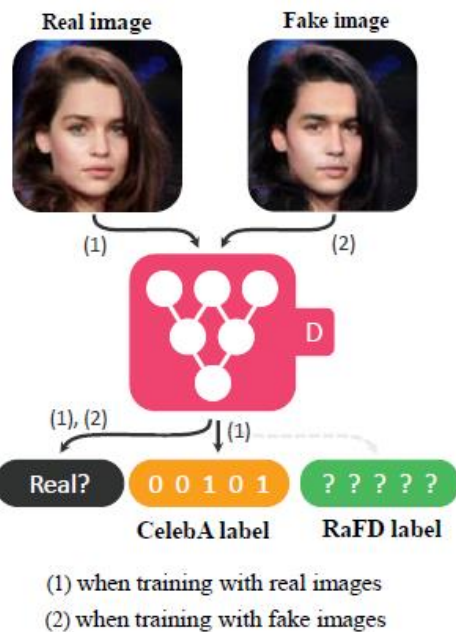
Mask vector로 해결

- 데이터셋이 여러 개일 경우 도메인 라벨 벡터

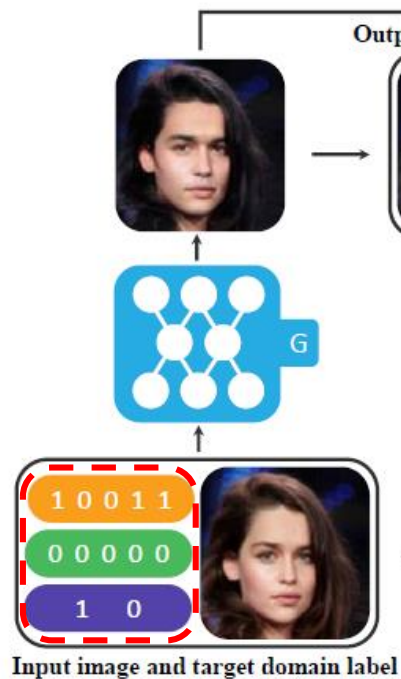




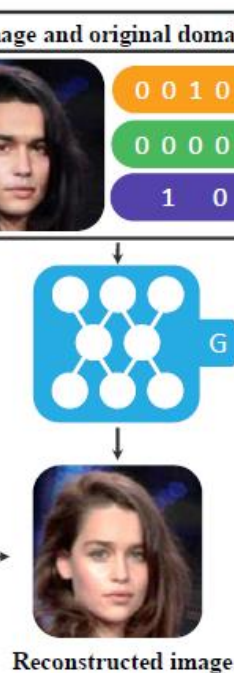
(a) Training the discriminator



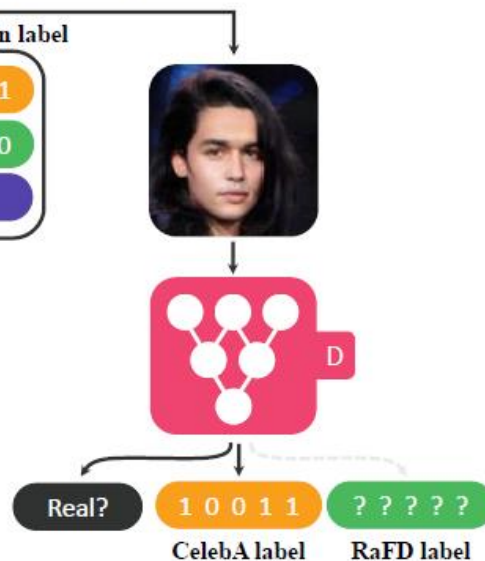
(b) Original-to-target domain



(c) Target-to-original domain



(d) Fooling the discriminator



Training with CelebA

CelebA label

Black / Blond / Brown / Male / Young

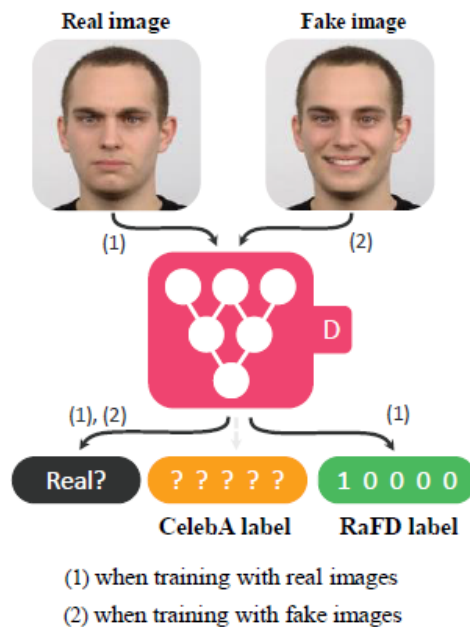
RaFD label

Angry / Fearful / Happy / Sad / Disgusted

Mask vector

CelebA / RaFD

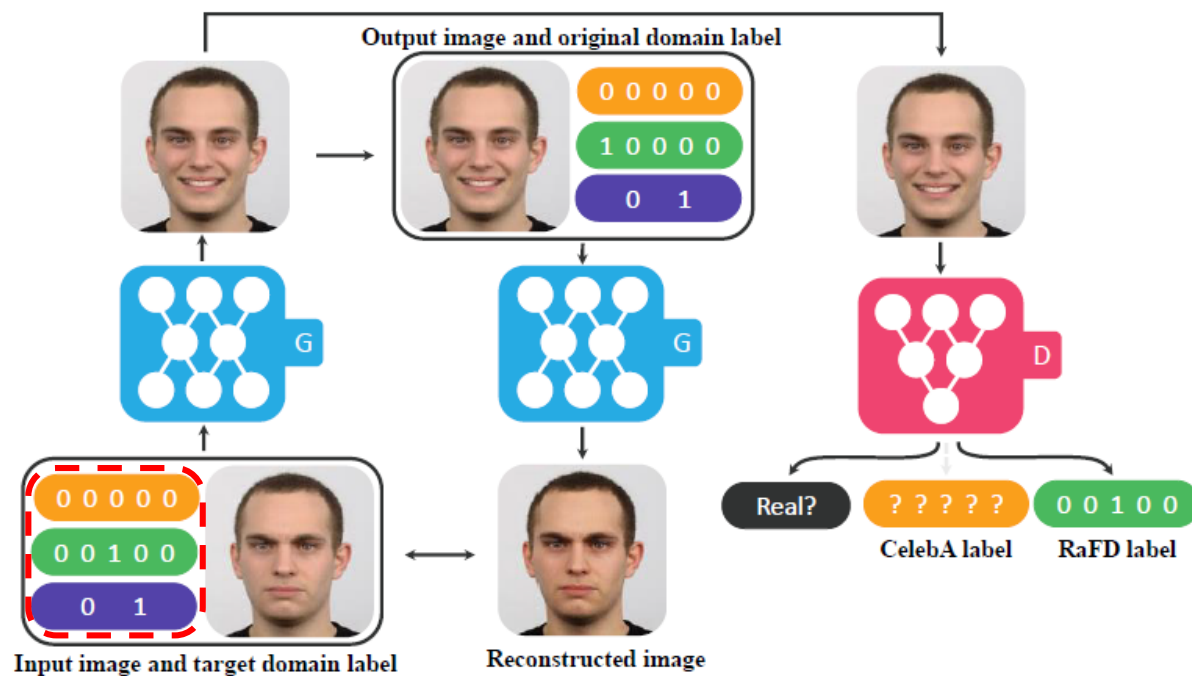
(e) Training the discriminator



(f) Original-to-target domain

(g) Target-to-original domain

(h) Fooling the discriminator



Experimental Results

Experimental Results

다른 baseline 모델에 비해 잘 만들어준다!

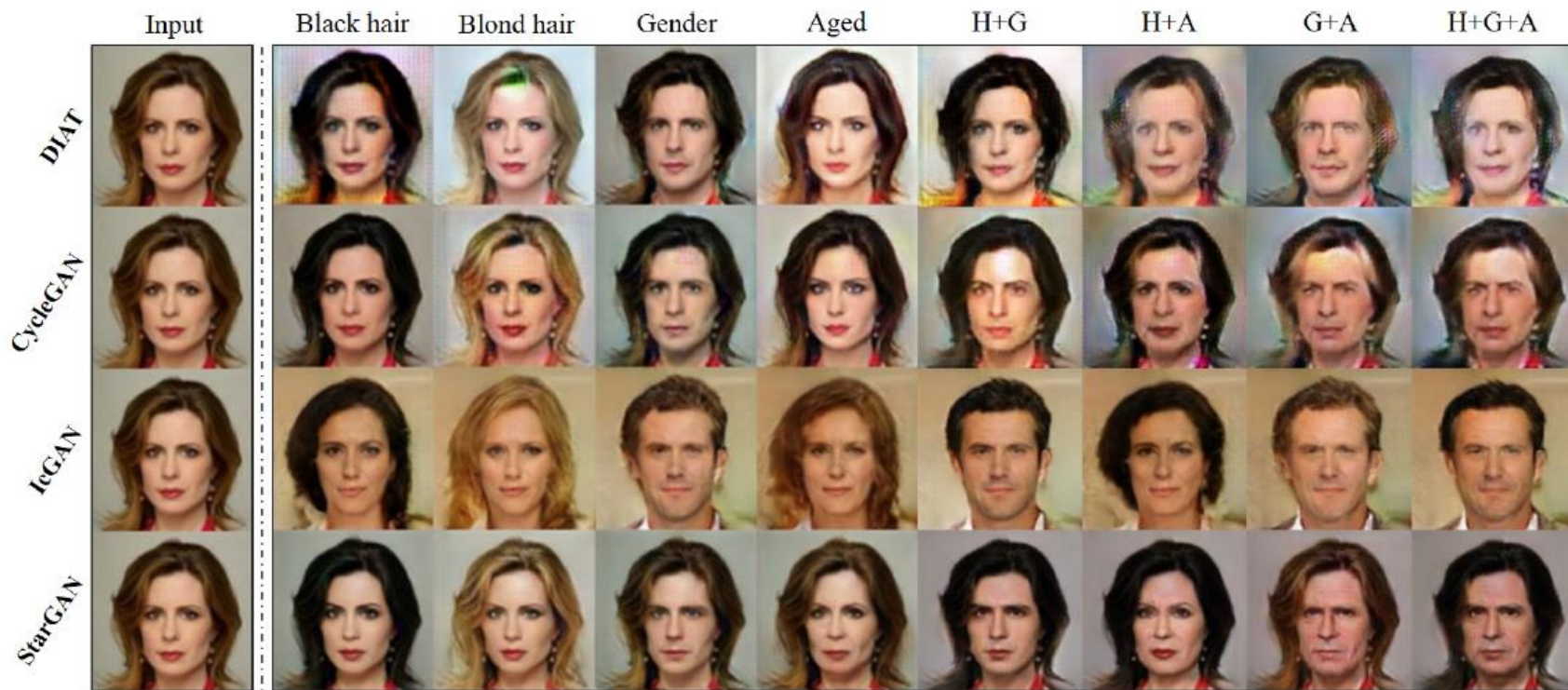


Figure 4. Facial attribute transfer results on the CelebA dataset. The first column shows the input image, next four columns show the single attribute transfer results, and rightmost columns show the multi-attribute transfer results. H: Hair color, G: Gender, A: Aged.

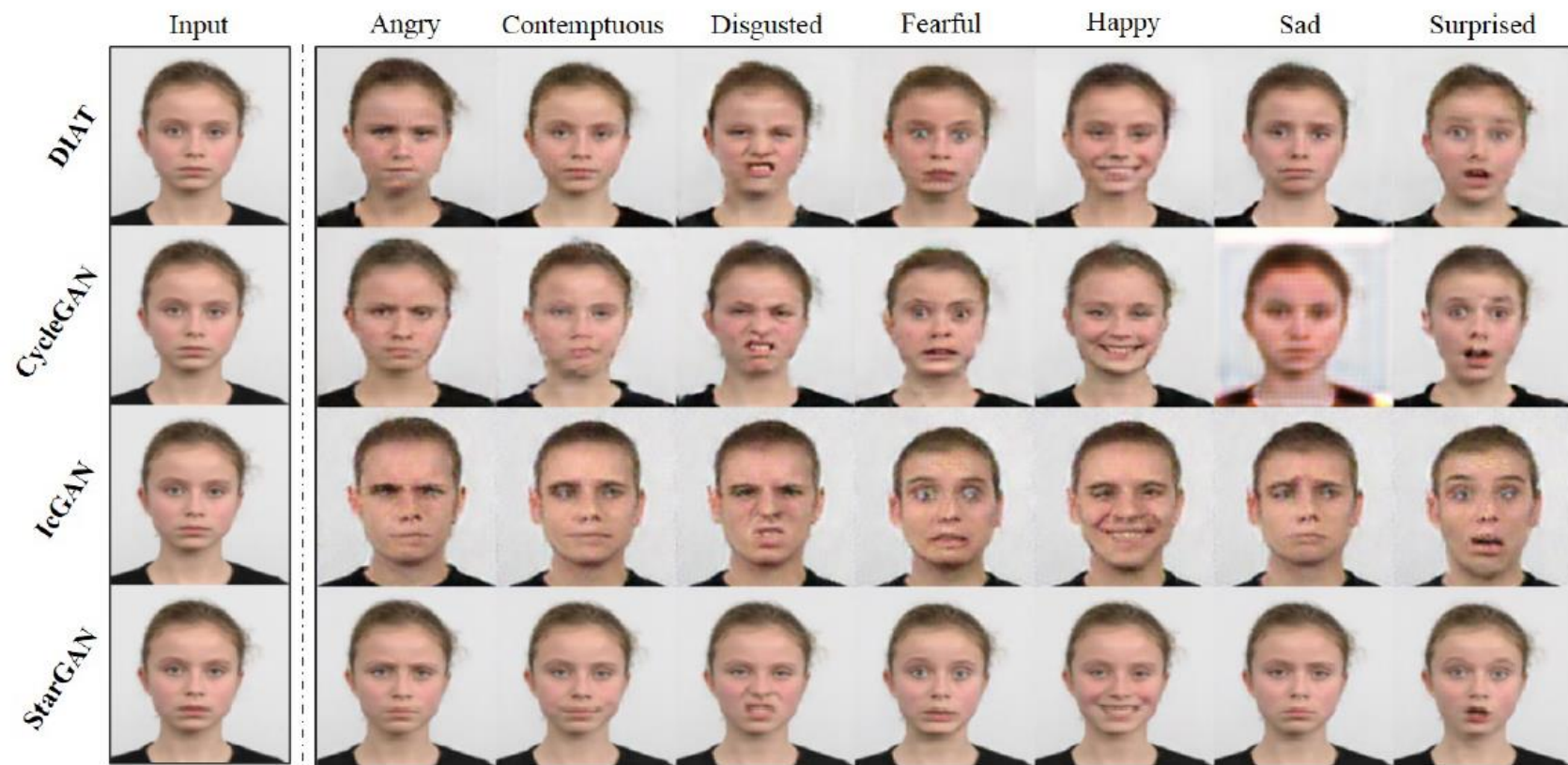


Figure 5. Facial expression synthesis results on the RaFD dataset.

Amazon Mechanical Turk (AMT)로 평가한 결과

Method	Hair color	Gender	Aged
DIAT	9.3%	31.4%	6.9%
CycleGAN	20.0%	16.6%	13.3%
IcGAN	4.5%	12.9%	9.2%
StarGAN	66.2%	39.1%	70.6%

Method	H+G	H+A	G+A	H+G+A
DIAT	20.4%	15.6%	18.7%	15.6%
CycleGAN	14.0%	12.0%	11.2%	11.9%
IcGAN	18.2%	10.9%	20.3%	20.3%
StarGAN	47.4%	61.5%	49.8%	52.2%

Experimental Results

다른 baseline 모델에 비해 잘 만들어준다!

그러나 주목할 만한 결과는 따로 있음!

“Scalability”

Method	Classification error	# of parameters
DIAT	4.10	52.6M × 7
CycleGAN	5.99	52.6M × 14
IcGAN	8.07	67.8M × 1
StarGAN	2.12	53.2M × 1
Real images	0.45	-

각 모델이 생성한 표정이 변환된 이미지들을 ResNet-18로 학습한 표정 분류기(정확도 99.55%)로 분류했을 때의 오차

각 모델의 학습에 사용된 **parameter** 수

Joint 학습의 효과

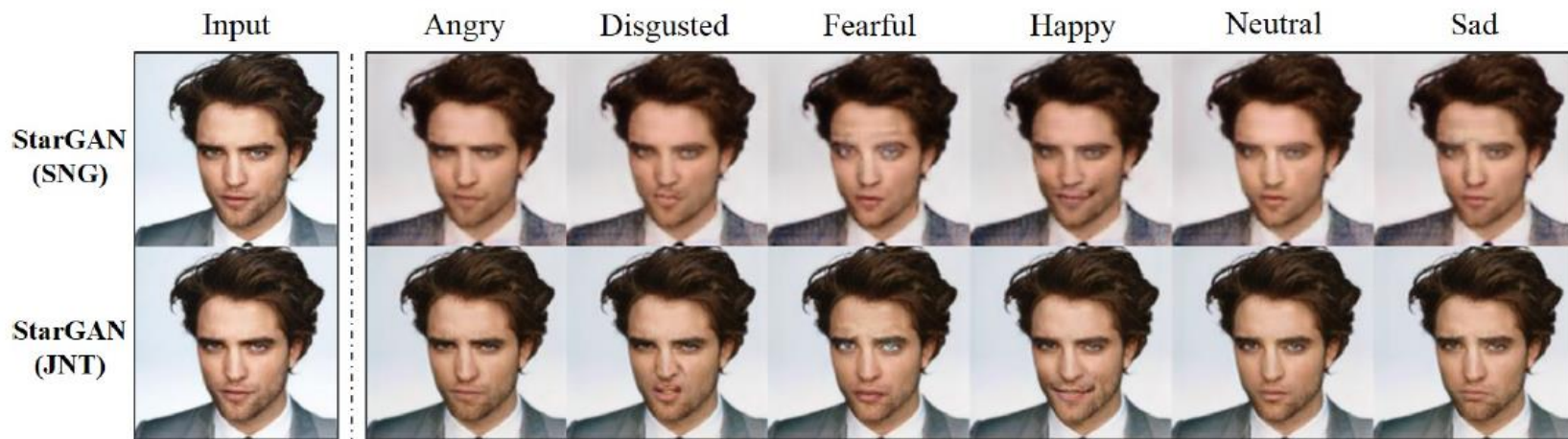
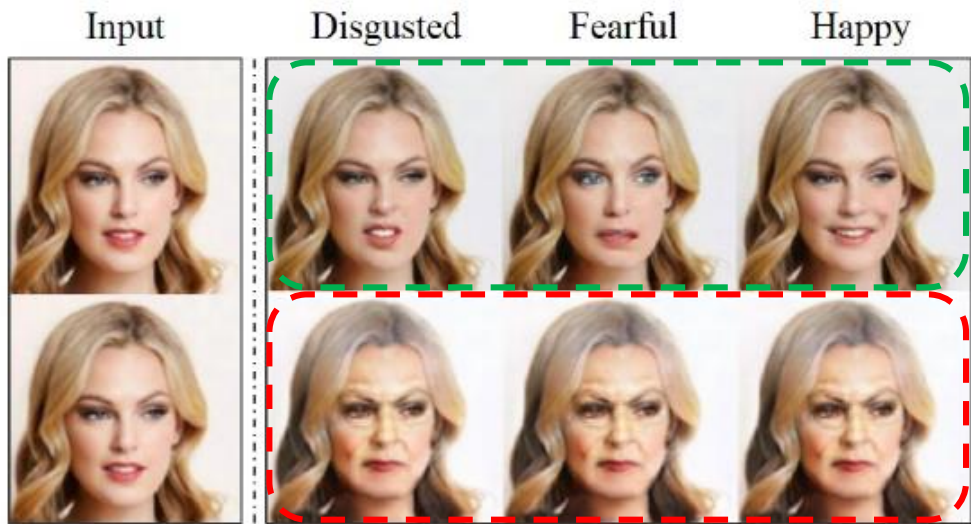


Figure 6. Facial expression synthesis results of StarGAN-SNG and StarGAN-JNT on CelebA dataset.

Mask vector 학습이 잘 되었는가?



올바른 mask vector일 경우
(0,1)

틀린 mask vector일 경우
(1,0)

Figure 7. Learned role of the mask vector. All images are generated by StarGAN-JNT. The first row shows the result of applying the proper mask vector, and the last row shows the result of applying the wrong mask vector.

내가 생각하는 StarGAN의 특이한 점!

감사합니다