
IoT·인공지능·빅데이터의 실무응용 연구 1

Various GANs and its Application : Face Aging

수리과학부 NCIA Lab

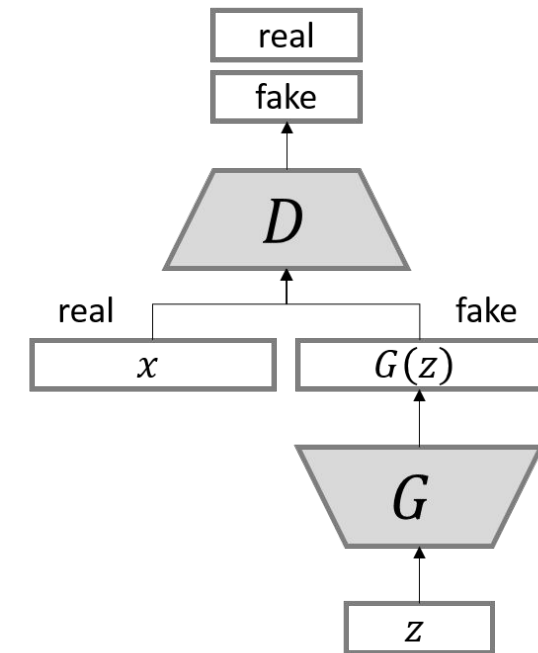
석사과정 송 기 웅



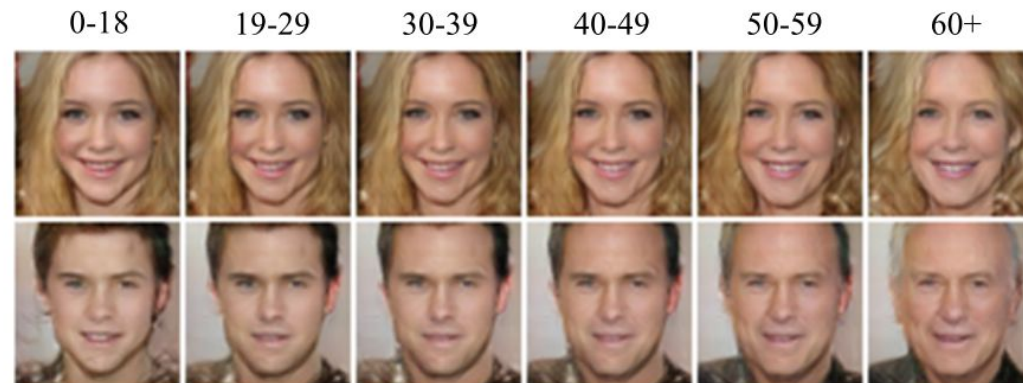
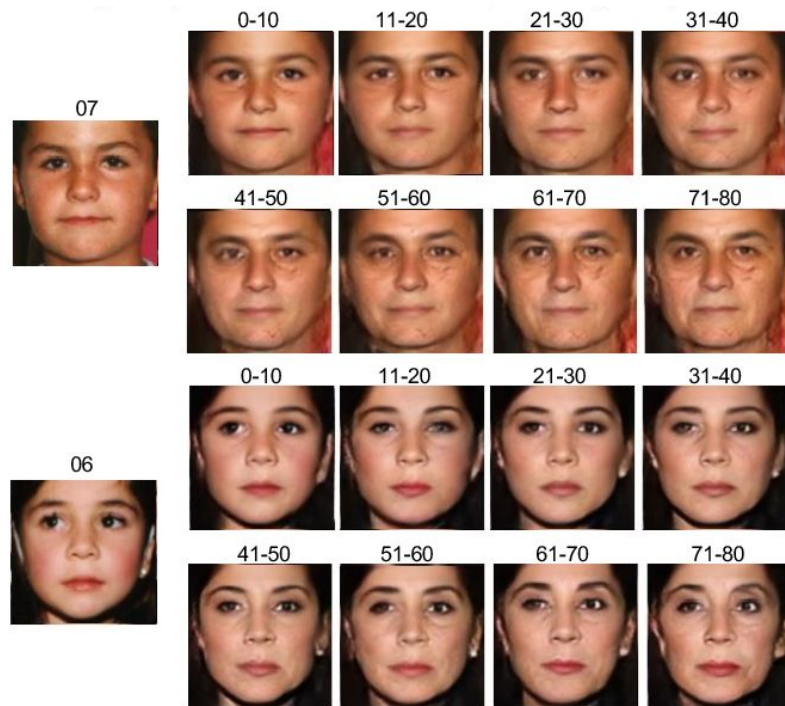
Various GANs

GAN

- Data의 분포를 학습하고 생성(샘플링)하기 위한 생성 모델
- Conditional Input의 활용 / 학습의 안정성 / 고해상도 Data의 생성 등을 위한 다양한 Architecture의 모델들이 존재 :
e.g. cGAN / WGAN / DCGAN / StyleGAN / CycleGAN
- Self-Supervised Learning 연구에 있어 활용되기 시작



Various GANs and its Application : Face Aging



Various GANs and its Application : Face Aging

1. Entertainment

2. Cross-age face recognition

e.g. 오래되어 시간 차이가 큰 신분증을 가진 사람의 신원 확인

3. Finding lost children

e.g. KIST PoliSketch

Face Aging with GANs

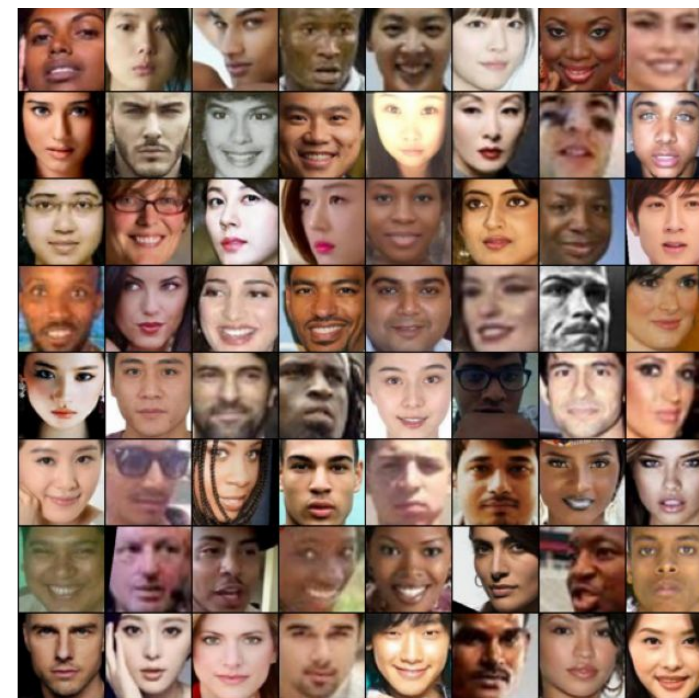
0. cGAN(Conditional GAN) based approach
1. CycleGAN based approach
2. StarGAN based approach
3. StarGAN based approach + Self-Attention

Dataset : UTKFace

“UTKFace - Large Scale Face Dataset Cropped”(100MB)

데이터셋을 활용하여 학습을 진행

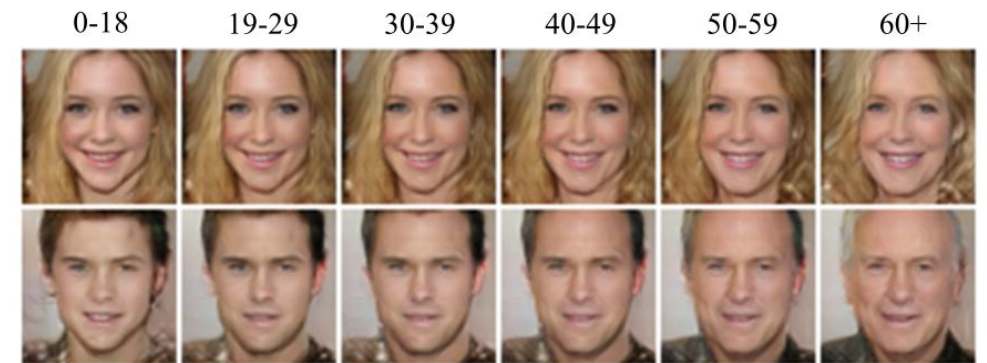
- 23000여장의 이미지 & 64 x 64 size의 작은 이미지로 다양한 학습을 시도
- 17 - 29세 / 30 - 39세 / 40 - 59세 / 60+세 의 4개의 그룹으로 나누어 학습을 진행



cGAN based approach

Conditional GAN(cGAN)

- 기존 GAN에 Conditional Input을 더하여 data를 생성하기 위한 모델
- Face Aging with Conditional Generative Adversarial Networks 라는 논문에서 이러한 cGAN을 활용하여 Face Aging을 수행하는 모델 제시

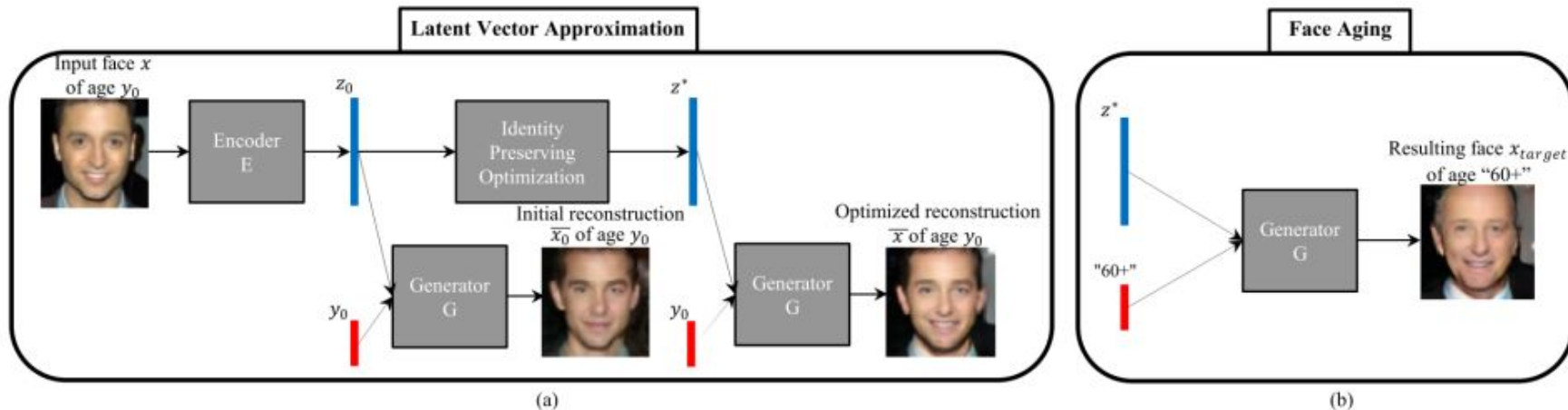


cGAN based approach

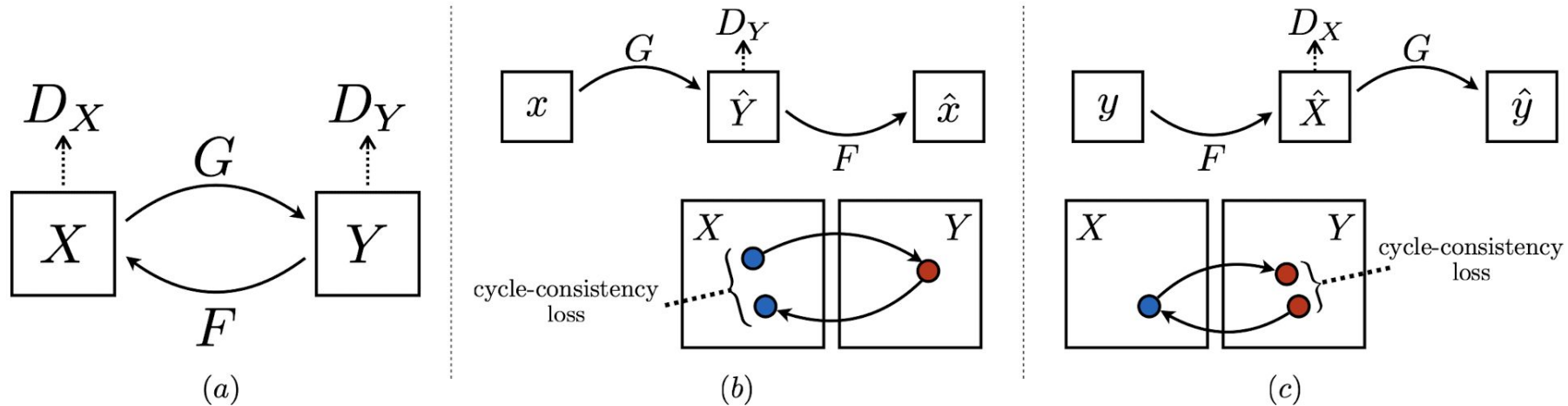
- Identity Preserving

Face Aging 과정에서 원래 사람의 얼굴을 보존하기 위한 Feature Optimization 부분이 필요

→ GAN 학습 이외에 별도의 Supervised Learning 모델을 필요로 하는 비효율성 발생



CycleGAN based approach



2개의 **Generator**를 활용하여 **Cyclic**하게 학습이 이루어지는 형태를 통해 **cGAN** 모델만을 활용할 때와 달리 **Supervised Learning** 기반의 추가 모델을 필요로 하지 않으면서도 **Face Aging**에 있어 **Identity**를 유지할 것이라 생각

CycleGAN based approach

Original Images



Transferred Images



Original Images



Transferred Images

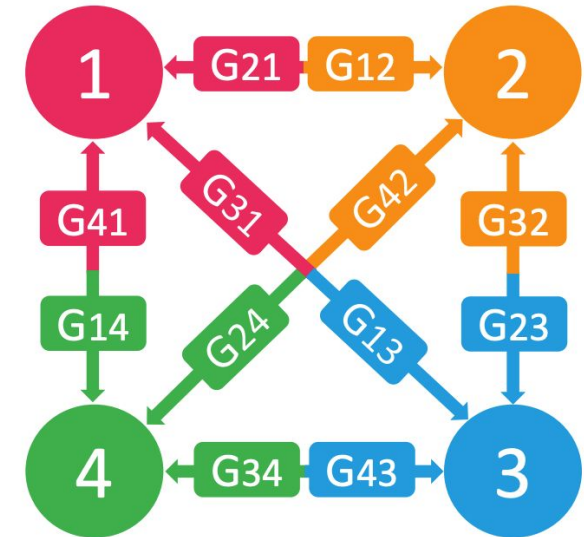


CycleGAN based approach

4개의 그룹의 경우 임의의 A 연령대에서 B 연령대로 얼굴을 바꾸기 위해서는 12개의 Generator를 학습해야하는 비효율 발생

→ cGAN과 같이 Conditional Input을 입력받는 형태를 추가하여 임의의 연령대로 translation을 학습하기 위한 Architecture 변형 생각 가능

(a) Cross-domain models

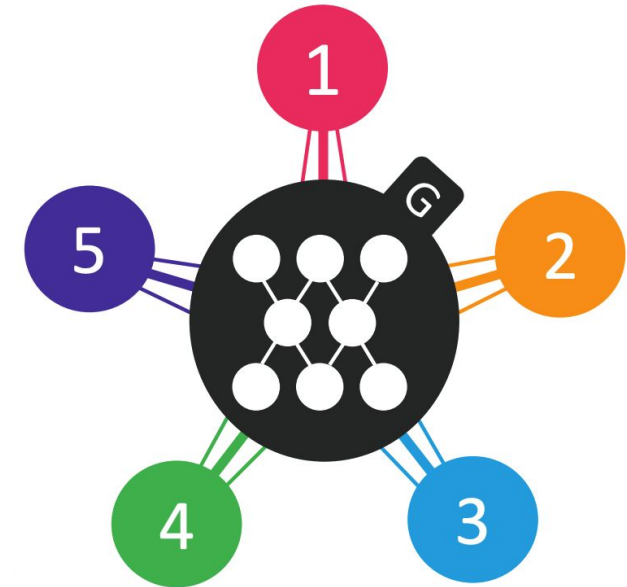


StarGAN based approach

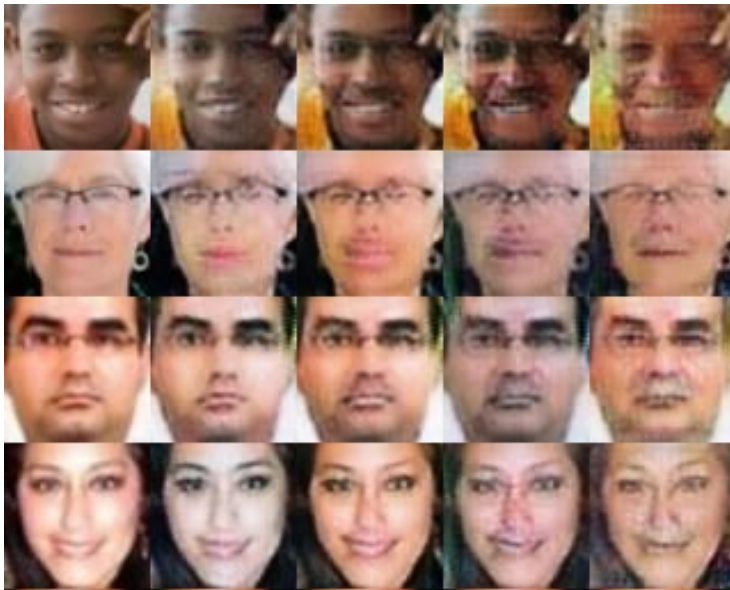
cGAN의 장점과 CycleGAN의 장점을 결합

- Face Aging에서 Identity를 유지하기 위한 별도의 Supervised Learning 모델이 필요없이 End to End로 학습이 가능
- Single Generator로부터 변화시키고자 하는 연령대를 Conditional Input으로 사용하여 Face Aging을 진행 가능

(b) StarGAN



StarGAN based approach



StarGAN based approach + Self-Attention

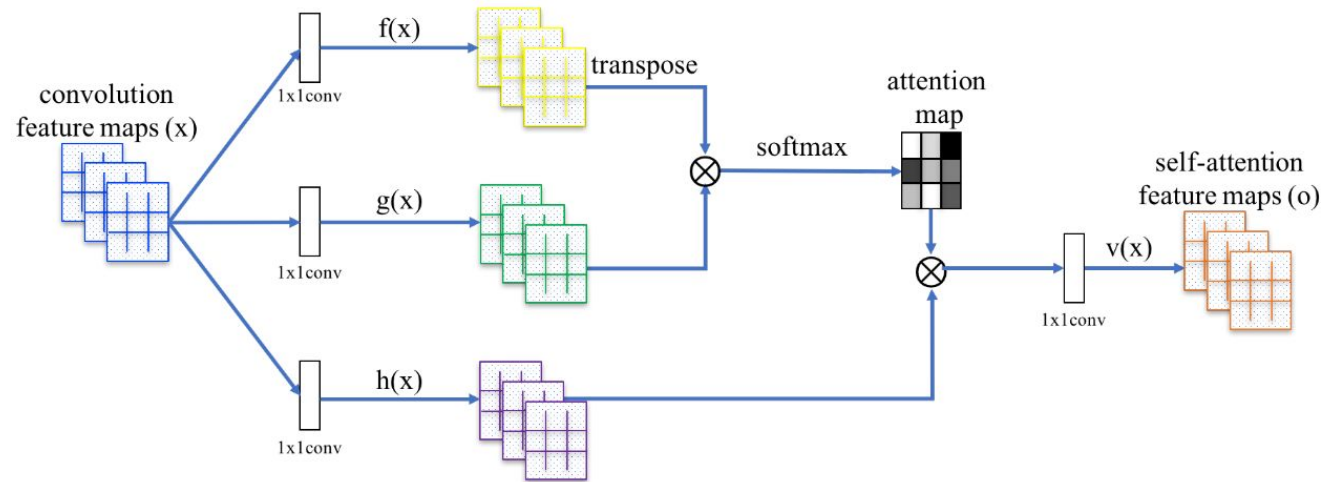


Figure 2. The proposed self-attention module for the SAGAN. The \otimes denotes matrix multiplication. The softmax operation is performed on each row.

Convolution을 통해 Local한 범위로 제한된 연산은 Global한 Image 구조 학습에 장애가 될 수 있어 효율적인 학습 모델 구성을 위해 SAGAN의 Self-Attention 모듈을 추가

Quantitative Evaluation : FID

FID(Fréchet Inception Distance)

- Inception Network를 사용하여 실제 이미지와 생성된 이미지의 Feature를 중간 Layer에서 계산을 하여 거리를 계산하는 평가 지표
- FID 값이 작을수록 실제 이미지와 생성된 이미지가 분포적으로 가까운 것을 의미하기에 GAN 모델의 좋은 성능을 의미

Quantitative Evaluation : FID

17 - 29세 연령대에서 60+ 이상 연령대로의 FaceAging에 대해 FID를 계산

→ UTKFace Crop : 64 x 64 size Image로 학습된 기준

	FID score
CycleGAN	232.6188
StarGAN	353.4051
StarGAN + Self - Attention	246.5845

Qualitative Evaluation

	FID score
CycleGAN	232.6188
StarGAN	353.4801
StarGAN + Self - Attention	246.5845

Original Images



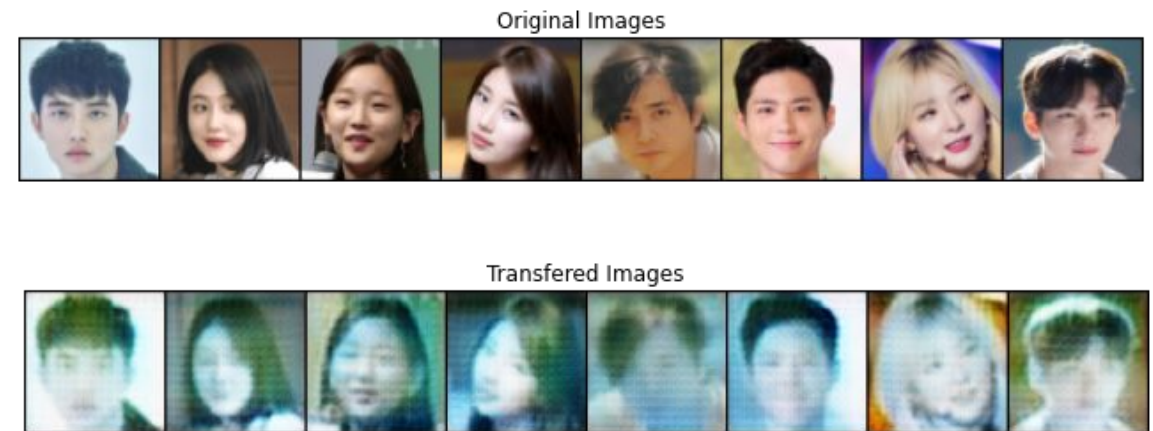
Transferred Images



CycleGAN

Qualitative Evaluation

	FID score
CycleGAN	232.6188
StarGAN	353.4801
StarGAN + Self - Attention	246.5845

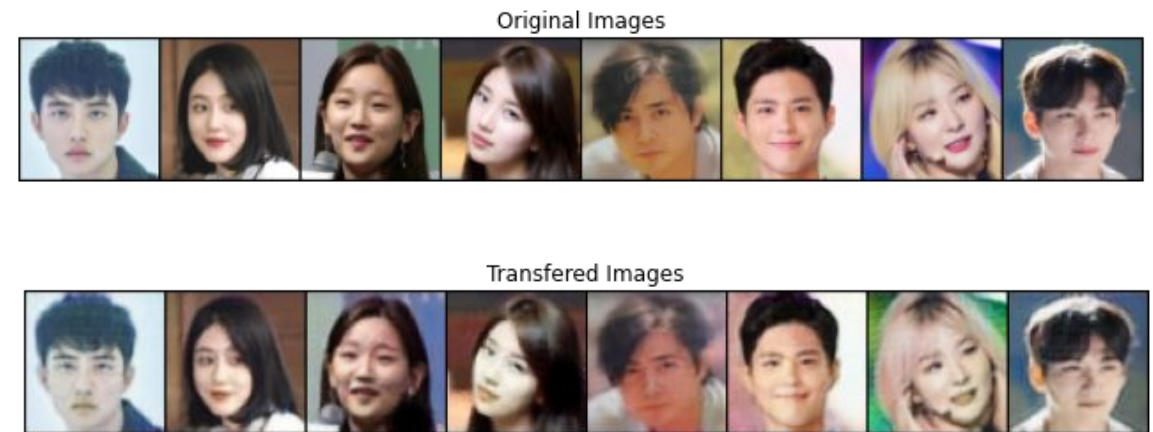


StarGAN

Generator와 Discriminator의 Residual-block만을 많이 쌓는 것만으로는 충분한 translation 성능을 보이지 못하였음

Qualitative Evaluation

	FID score
CycleGAN	232.6188
StarGAN	353.4801
StarGAN + Self - Attention	246.5845



StarGAN + Self-Attention

Generator가 input을 그대로 복제하는 translation만으로도 Discriminator를 충분히 속이는 Mode Collapse가 발생하였음

Qualitative Evaluation

Original Images



Transferred Images



CycleGAN / FID : 232.6188

Original Images



Transferred Images



StarGAN + Self-Attention / FID : 246.5845

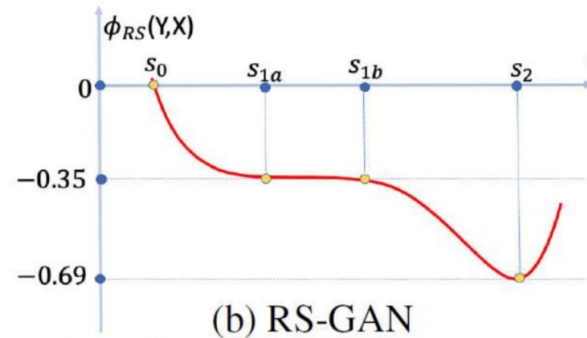
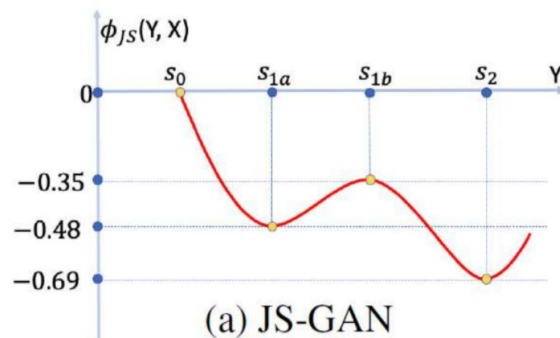
64 x 64 크기의 Input Image로는 Discriminator 4개의 연령대를 유의미하게 구분하는 것에 성능의 제약이 있다고 판단됨

Appendix

How to explain mode collapse of GAN mathematically?

→ “Towards a Better Global Loss Landscape of GANs” at NeurIPS 2020

-
- $s_0: |\{x_1, x_2\} \cap \{y_1, y_2\}| = 0$: “no-recovery”
 - $s_{1a}: y_1 = y_2 \in \{x_1, x_2\}$: “mode collapse”
 - $s_{1b}: |\{x_1, x_2\} \cap \{y_1, y_2\}| = 1$: “mode dropping”
 - $s_2: \{x_1, x_2\} = \{y_1, y_2\}$: “perfect-recovery”



Appendix

Original Images



Transferred Images



StarGAN + Self-Attention / FID : 246.5845
(64 x 64 size Input Image)

Original Images



Transferred Images



StarGAN + Self-Attention / FID : 228.9953
(128 x 128 size Input Image with Hinge loss)

Appendix

Original Images



Transferred Images



CycleGAN / FID : 232.6188
(64 x 64 size Input Image)

Original Images



Transferred Images



StarGAN + Self-Attention / FID : 228.9953
(128 x 128 size Input Image with Hinge loss)

Appendix

Original Images



Transferred Images



StarGAN + Self-Attention / FID : 228.9954
(128 x 128 size Input Image with Hinge loss)

Original Images



Transferred Images



StarGAN + Self-Attention / FID : 240.5232
(128 x 128 size Input Image with Hinge loss)

Reference

- [1] Antipov, G., Baccouche, M., & Dugelay, J.-L. (2017). Face Aging With Conditional Generative Adversarial Networks.
- [2] Choi, Y., Choi, M., Kim, M., Kim, J. H. S., & Choo, J. (2018). StarGAN : Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation.
- [3] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... Bengio, Y. (2014). Generative Adversarial Nets.
- [4] Sun, R., Fang, T., & Schwing, A. (2020). Towards a better global loss landscape of GANs. ArXiv, (NeurIPS).
- [5] Zhang, H., Goodfellow, I., Metaxas, D., & Odena, A. (2018). Self-Attention Generative Adversarial Networks.
- [6] Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. Proceedings of the IEEE International Conference on Computer Vision, 2017-Octob, 2242–2251.

| Q & A

