[Deep Learning Seminar]
Season 2 : Deep Learning Paper Review

Learning to Discover Cross-Domain Relations with Generative Adversarial Networks

By Taeksoo Kim et al.

Hanyang University

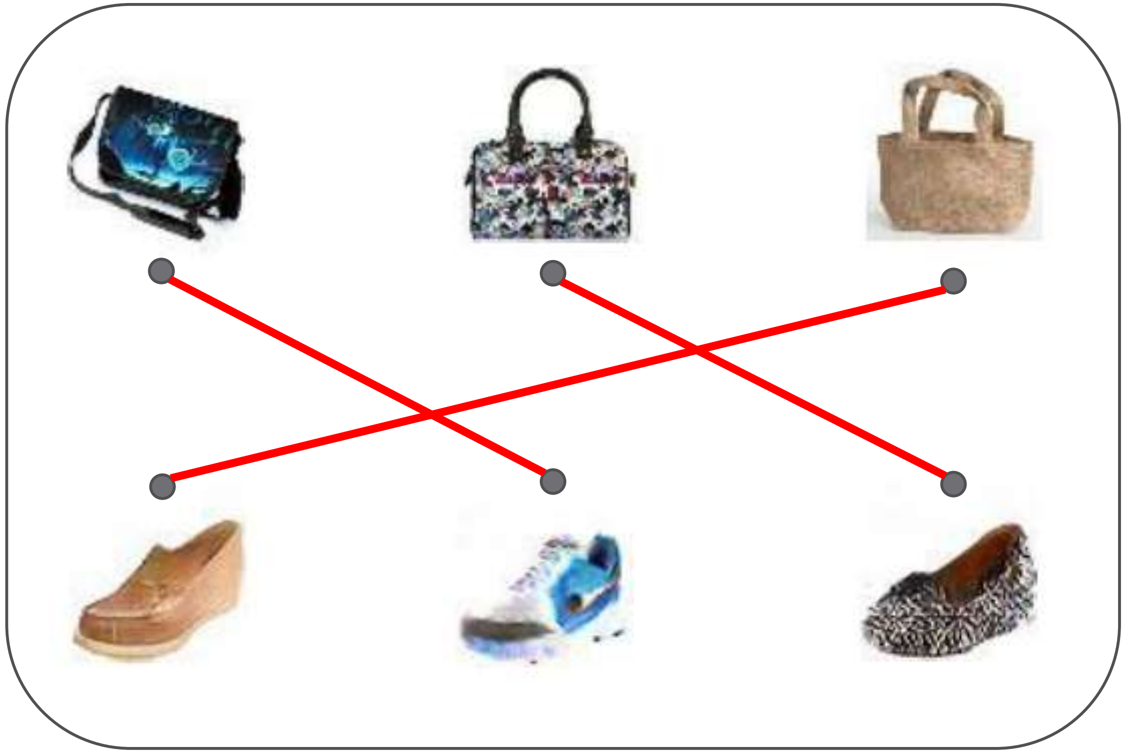
Artificial Intelligence Lab.

Gunhee Cho

Contents

Introduction

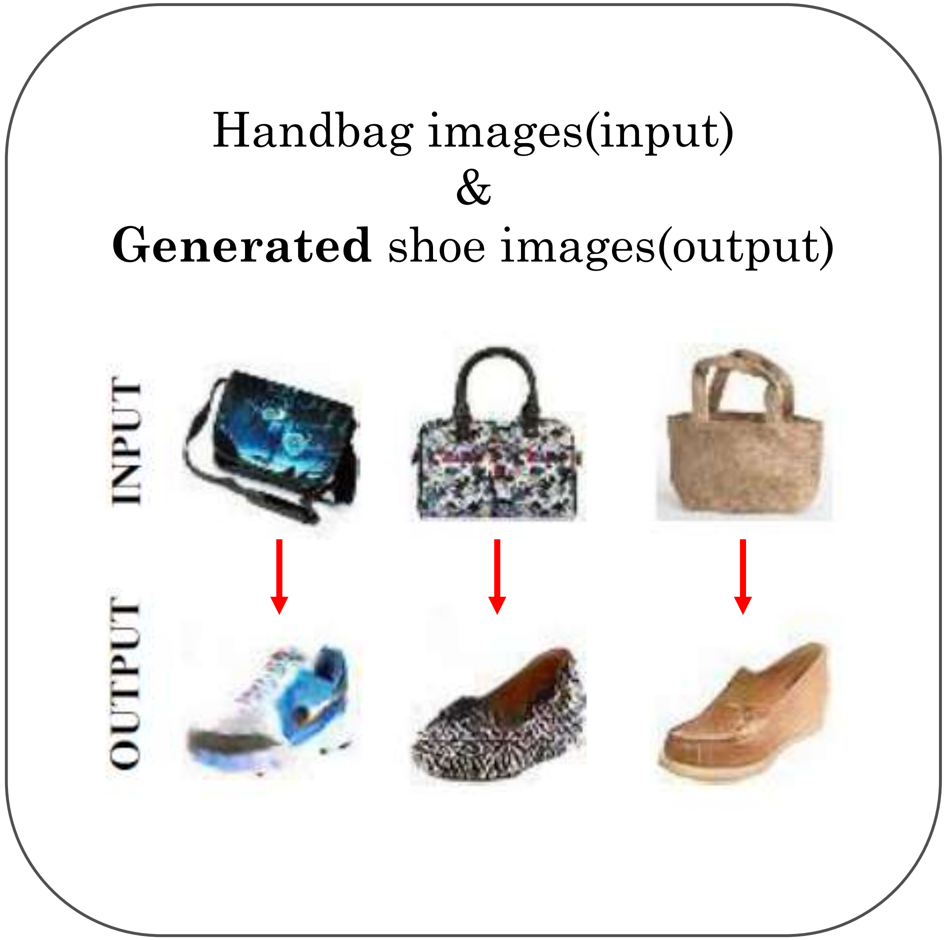
- "비슷한 스타일(Similar style)"을 이해하는 것은 사람에게는 아주 쉬운 일.
- 즉, 서로 다른 도메인의 데이터에서 관련성 (**Cross-domain relation**)을 찾아내는 것을 사람은 쉽게 할 수 있음.
- 오른쪽 예시
  - 공통점 : 비슷한 스타일/질감
  - 차이점 : 도메인(가방/신발)



Introduction

- 서로 다른 도메인의 두 이미지의 연관성을 찾는 것을 학습시키려면 어떻게 해야 할까?

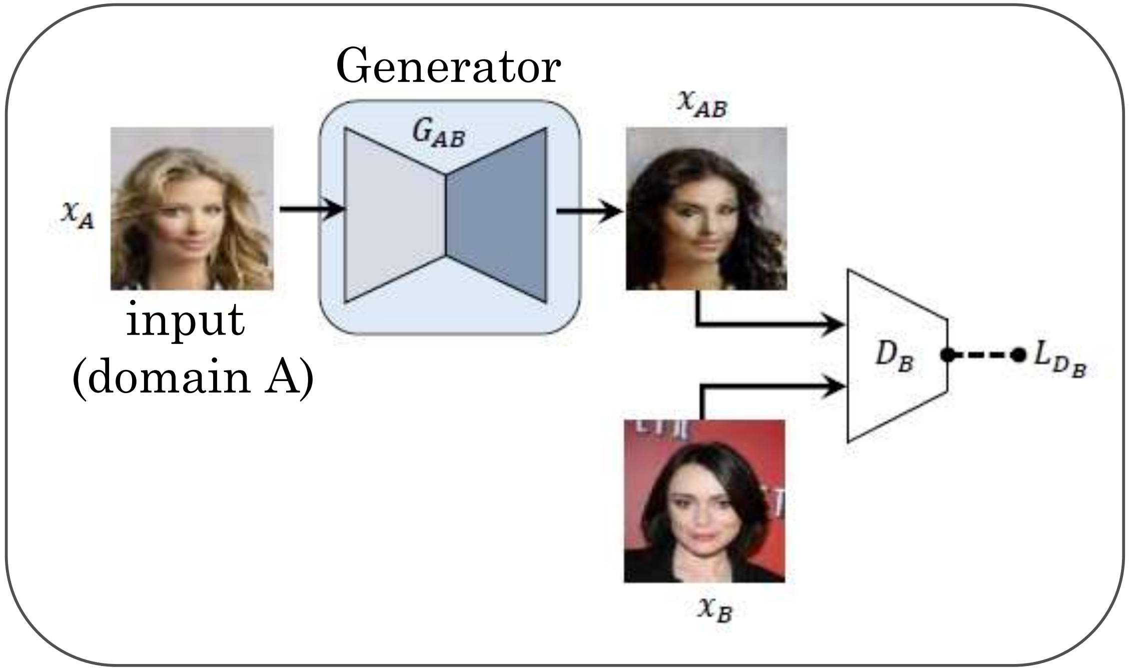
→이 논문에서는 **"Cross domain relation discovery"** 문제를 해결하는 것이 조건부 이미지 생성 문제(conditional image generation)를 해결하는 것과 같다고 생각함.



Model - Baseline models

- Standard GAN
  - Relation Discovery Task 를 위해 가장 먼저 시도한 모델.
  - 기존 GAN 은 이미지를 생성해주는 Generator 의 input으로 Gaussian noise를 사용하지만, 이 논문에서는 GAN의 input으로 이미지 자체를 사용.
  - 이러한 모델을 이용하면 오직 한쪽 방향 (domain A → domain B)으로만 학습이 가능함.

$$x_{AB} = \boldsymbol{G}_{AB}(x_A)$$



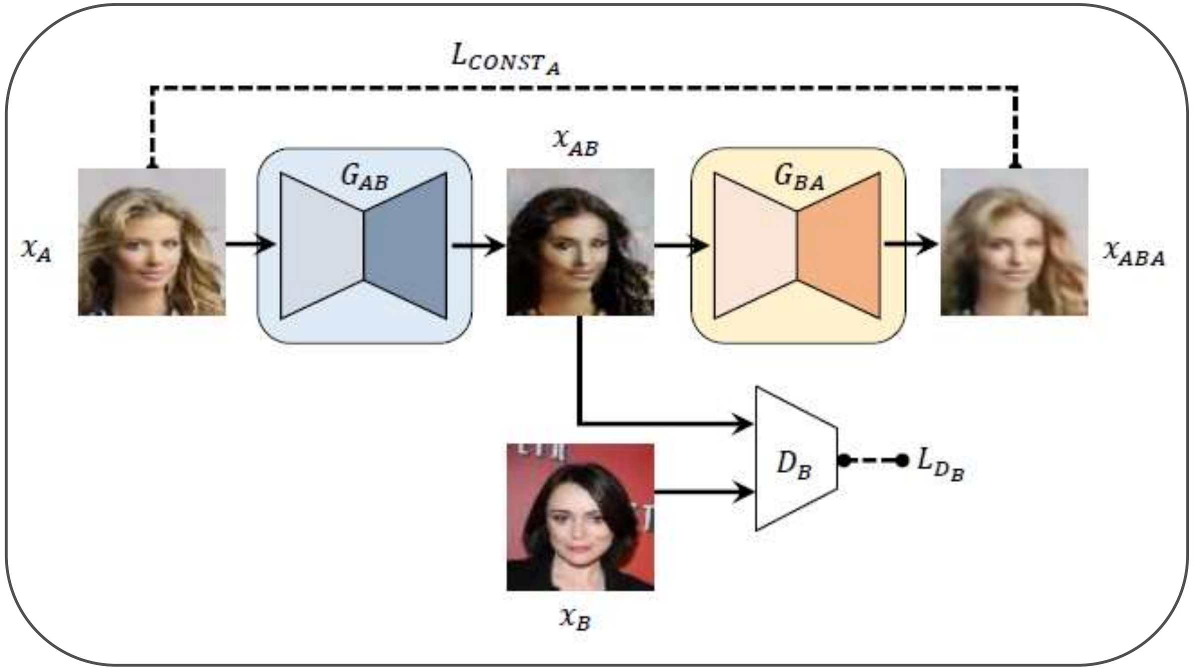
Model - Baseline models

- GAN with reconstruction loss
  - Standard GAN 모델에 Reconstruction 단계 ($\boldsymbol{G}_{BA}$)를 추가해준 모델.
    - ✓ L_{GAN_B} : Standard GAN의 generator Loss
    - ✓ L_{CONST_A} : Reconstruction Loss
  - Standard GAN Loss와 Reconstruction Loss를 동시에 최적화해주는 식으로 학습 진행.
    - ✓ $L_{G_{AB}} = L_{GAN_B} + L_{CONST_A}$

How much similar?

How much realistic?

$$
\begin{aligned}
x_{AB} &= \boldsymbol{G}_{AB}(x_A) \\
x_{ABA} &= \boldsymbol{G}_{BA}(x_{AB}) = \boldsymbol{G}_{BA} \circ \boldsymbol{G}_{AB}(x_A) \\
L_{CONST_A} &= d(\boldsymbol{G}_{BA} \circ \boldsymbol{G}_{AB}(x_A), x_A) \\
L_{GAN_B} &= -\mathbb{E}_{x_A \sim P_A}[\log \boldsymbol{D}_B(\boldsymbol{G}_{AB}(x_A))]
\end{aligned}
$$

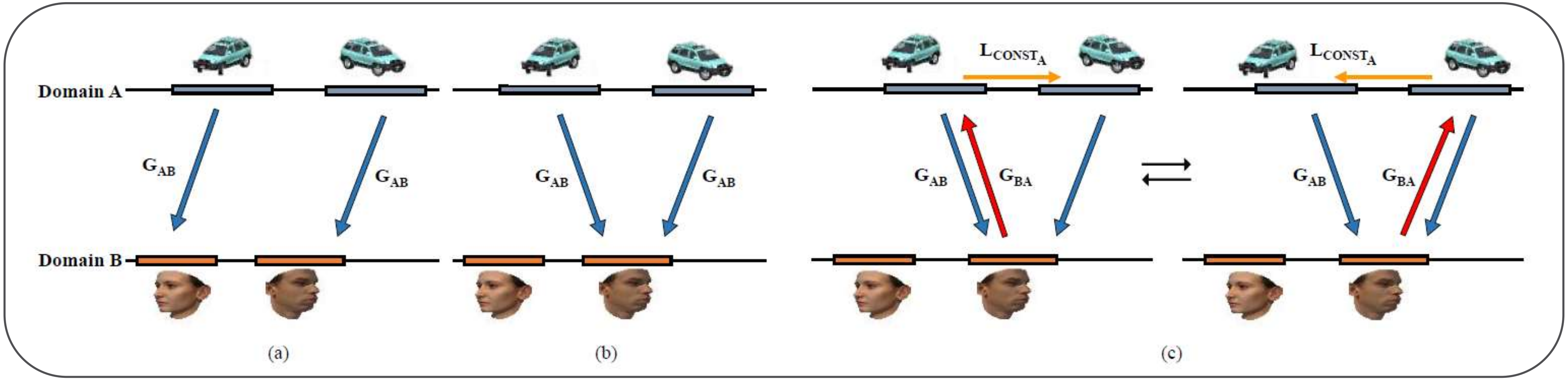


Model - Baseline models

- 한계점 : 한 방향(injection, not bijection)으로만 학습하기 때문에 Cross-Domain relation이 보장되지 않음.

  → **Mode collapse 현상**

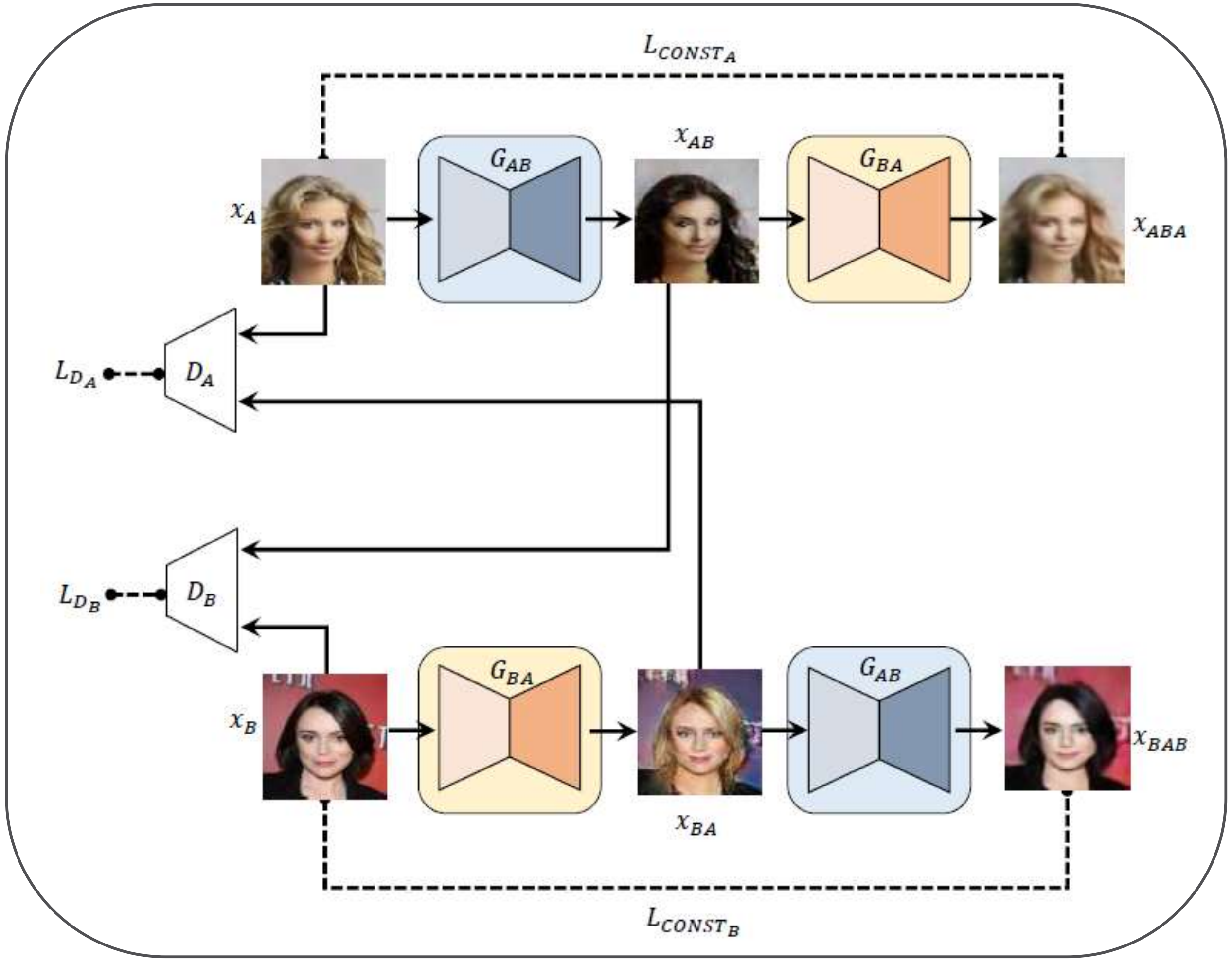
(a) ideal mapping
(b) GAN model failure
(c) GAN with reconstruction model failure



Model - DiscoGAN

- Discovery GAN
  - GAN with reconstruction 모델 2개를 결합한 형태의 모델.
  - 2개의 신경망이 동시에 학습되며 파라미터를 공유하기 때문에 bijective mapping 가능.
  - 즉, **Cross-domain relation discover**.
  - $L_G = L_{G_{AB}} + L_{G_{BA}}$

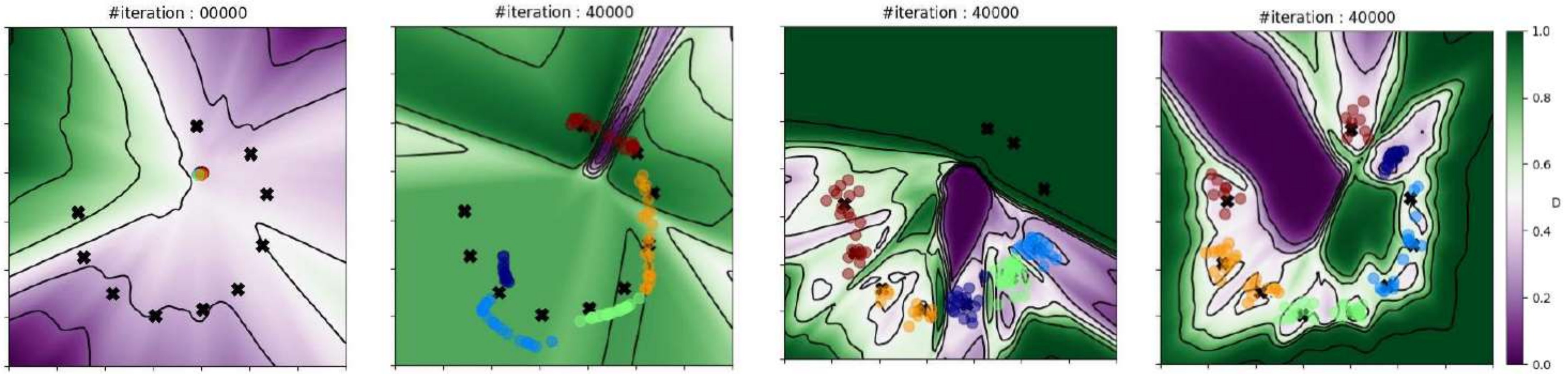
    $= L_{GAN_B} + L_{CONST_A} + L_{GAN_A} + L_{CONST_B}$
  - $L_D = L_{D_A} + L_{D_B}$



Experiment – Toy experiment

- Background color : discriminator의 output
- ‘x’ 마크 : B 도메인의 특정 모드(mode)
- color 점 : A 도메인에서 B 도메인으로 매핑한 결과

(a) initial state
(b) standard GAN
(c) GAN with reconstruction
(d) DiscoGAN

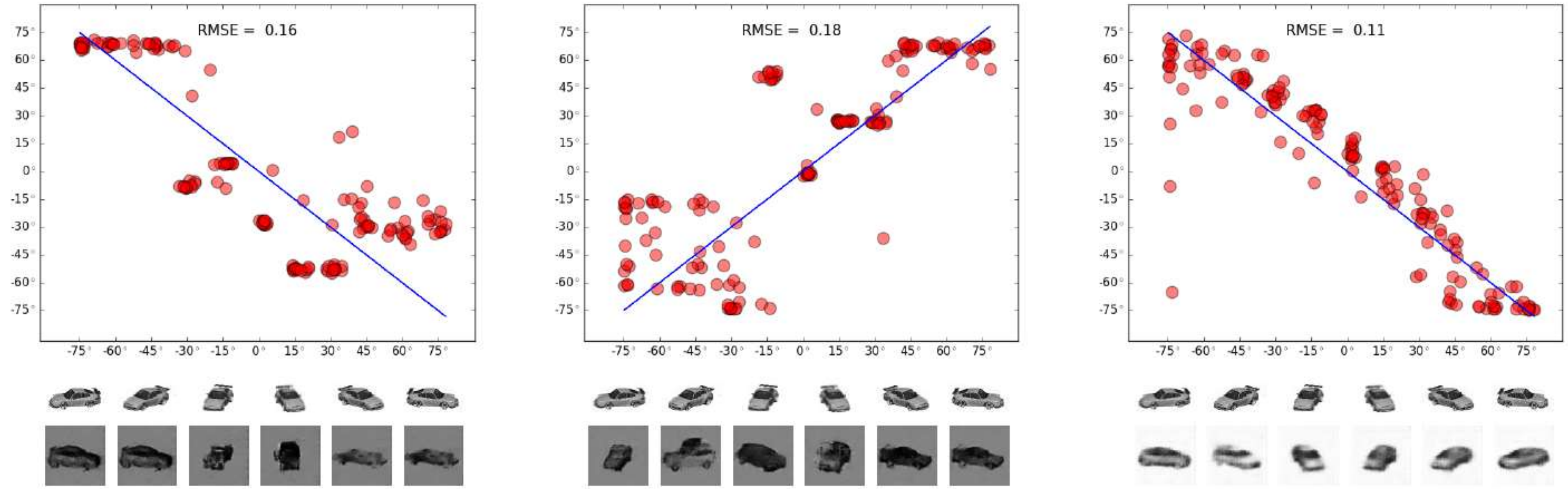


(a) (b) (c) (d)

Experiment – Real domain experiment

- Mode collapse 현상 해결
  - Car to Car translation experiment

(a) standard GAN
(b) GAN with reconstruction
(c) DiscoGAN

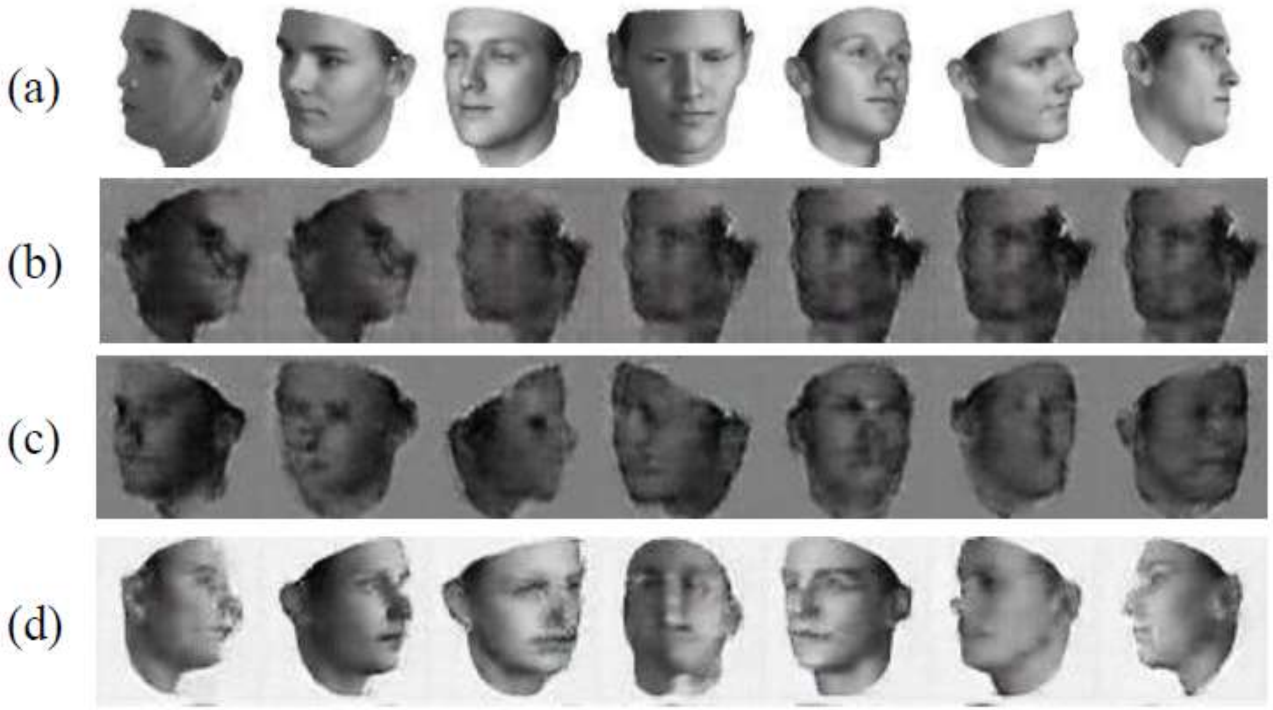


(a) (b) (c)

Correlation
가장 높음
=
Mode collapse
현상 해결

Experiment – Real domain experiment

- Mode collapse 현상 해결
  - Face to Face translation experiment



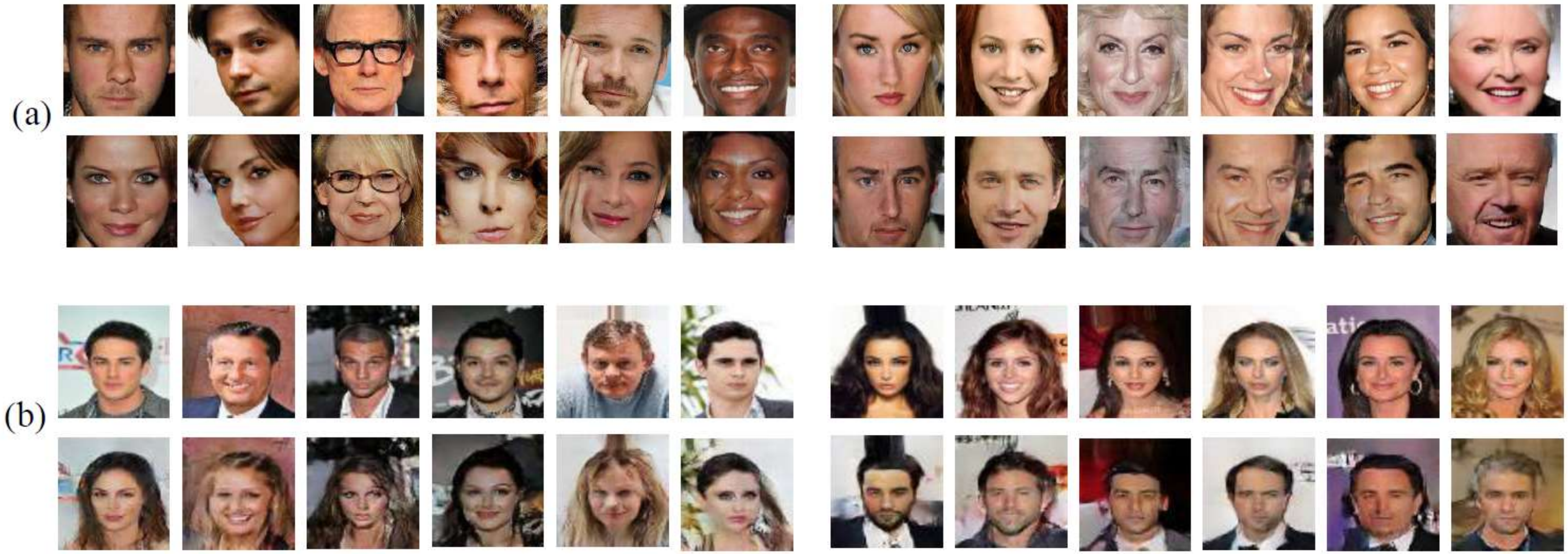
(a) input face image
(b) standard GAN
(c) GAN with reconstruction
(d) DiscoGAN

Mode collapse 현상 해결

Experiment – Real domain experiment

- 도메인 간 거의 모든 feature 공유
  - Face conversion

(a) gender translation
(b) gender translation with background



Experiment – Real domain experiment

- 도메인 간 거의 모든 feature 공유
  - Face conversion

(c) hair color translation
(d) eyeglasses translation

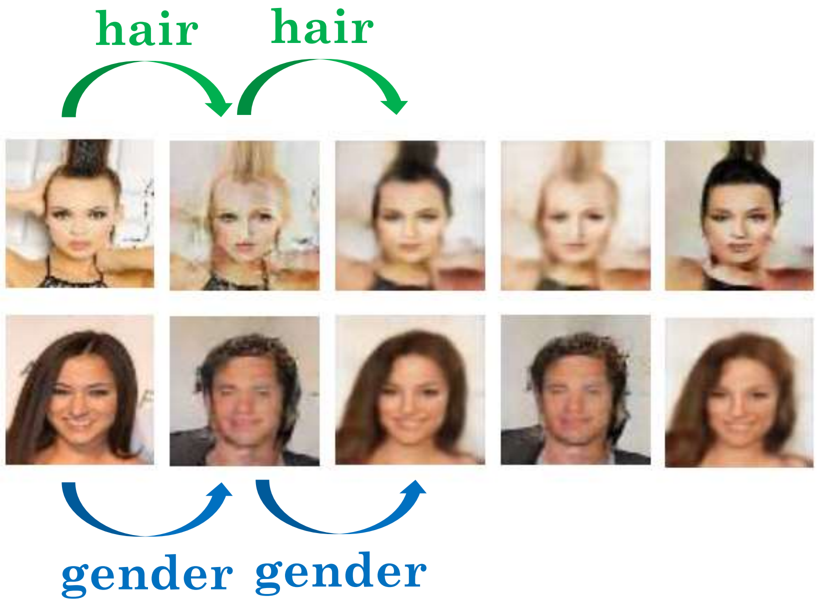
Experiment – Real domain experiment

- 도메인 간 거의 모든 feature 공유
  - Face conversion

(e) sequence of conversion of gender and hair
(f) repeatedly applying same conversions

(e)

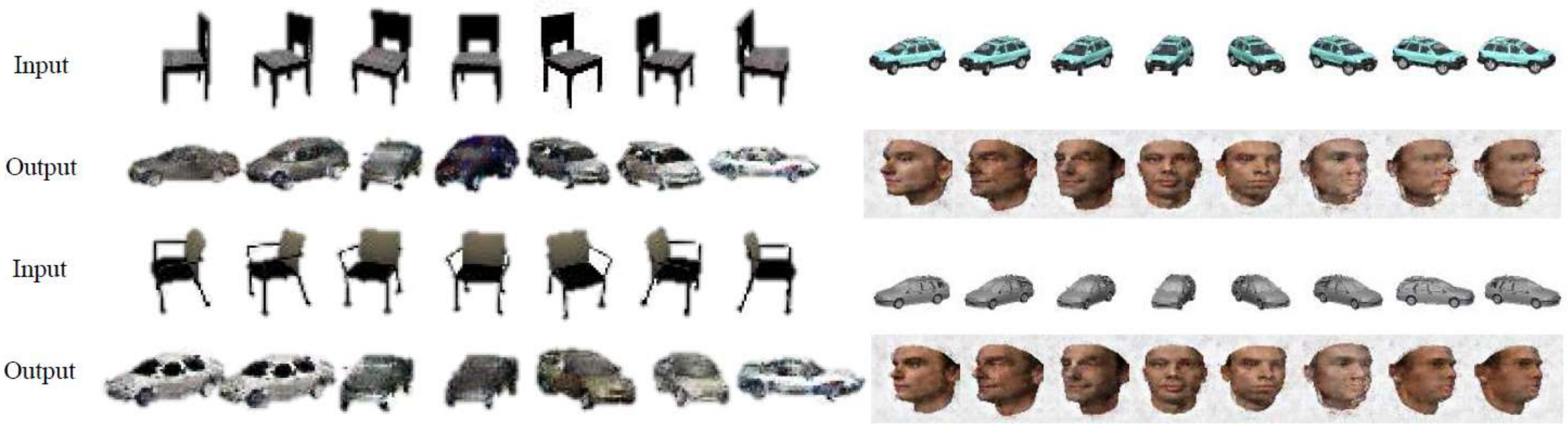
(f)



Experiment – Real domain experiment

- 도메인 간 feature 1개만 공유
  - Chair to Car
  - Car to Face

"바라보는 방향" 이라는 feature 하나만 같다.



(a) Chair to Car

(b) Car to Face

Experiment – Real domain experiment

- 1-to-N problem
  - Edge-to-Photos

(a) sketches of bags to colored image
(b) sketches of shoes to colored image
(c) colored image to sketches

Experiment – Real domain experiment

- 도메인 간 공유된 explicit feature 없음(visually very different)
  - Handbag to Shoes
  - Shoes to Handbag

(b) Handbag images (input) & **Generated** shoe images (output)

(c) Shoe images (input) & **Generated** handbag images (output)

Conclusion

- 논문이 소개한 것
  - **A learning method to discover cross-domain relations** with a generative adversarial network called DiscoGAN.
  - Without any explicit pair labels
  - Learn to relate datasets from very different domains
- 결론
  - DiscoGAN은 스타일 변화된 high-quality 이미지를 생성해낼 수 있다.
- 코드 예시
  - SK T-Brain official : https://github.com/SKTBrain/DiscoGAN
  - carpedm20(김태훈) : https://github.com/carpedm20/DiscoGAN-pytorch

Thank you