



덕업일치라고 할 뻔

웹툰 캐릭터 가상 캐스팅

16기 임채명 윤지현 17기 이서연 황우현

CONTENTS

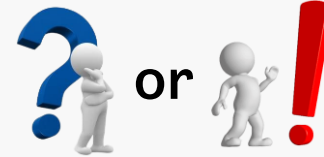
- 0 배경
- 1 목표 및 과정 요약
- 2 데이터 구축
- 3 모델링
- 4 한계

BANG

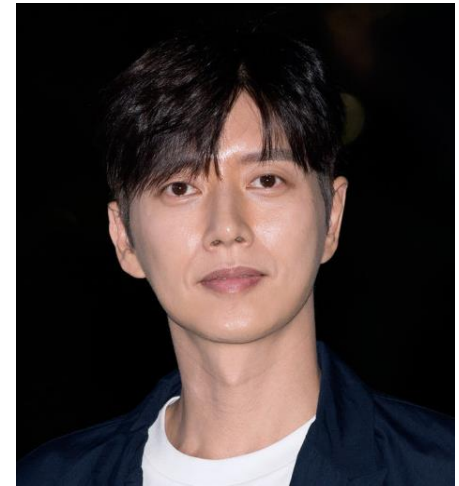
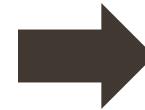


0 배경

웹툰 원작의 드라마 캐스팅 싱크로율에 큰 관심 ->



1 목표

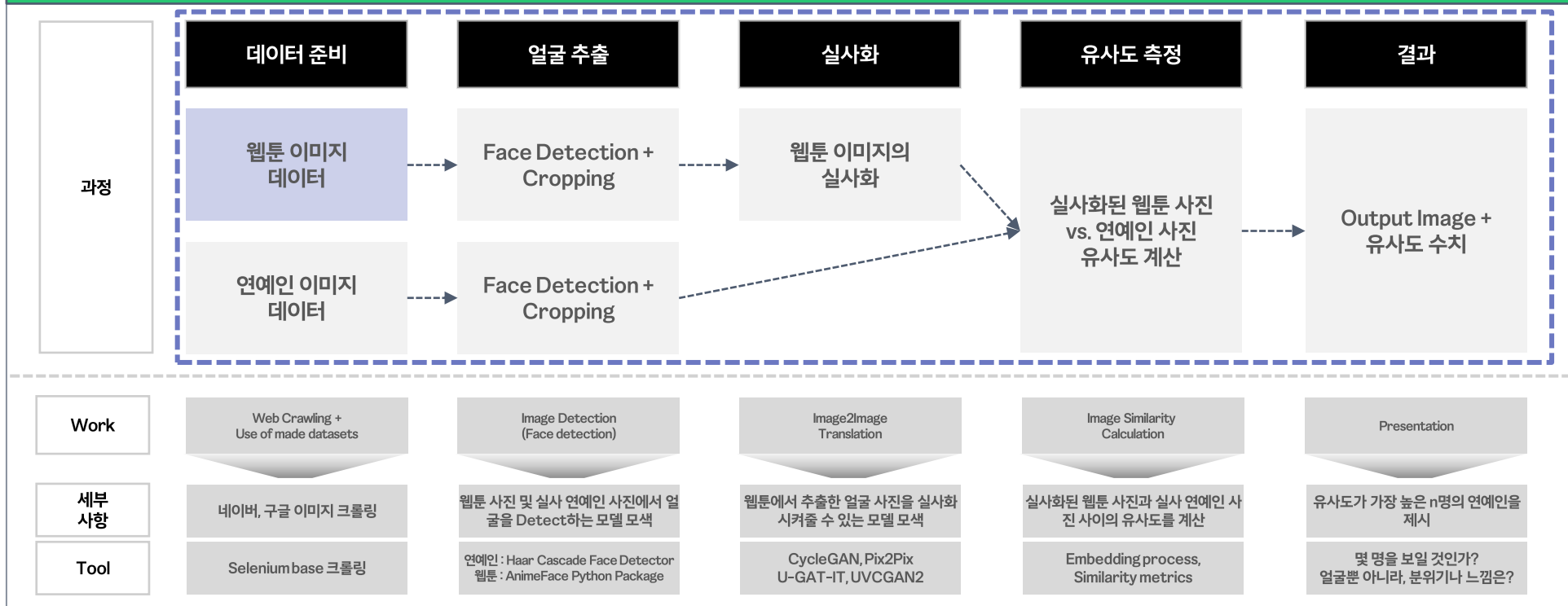


웹툰 캐릭터를 실제 인물화하여 유사한 연예인을 캐스팅 가능하게 한다.

1 과정 요약

Input

웹툰 가상캐스팅 - 웹툰을 드라마/영화化 할 때, 어떤 배우가 해당 역할에 가장 잘 어울릴까?



웹툰 캐릭터를 실제 인물화하여 유사한 연예인을 캐스팅 가능하게 한다.

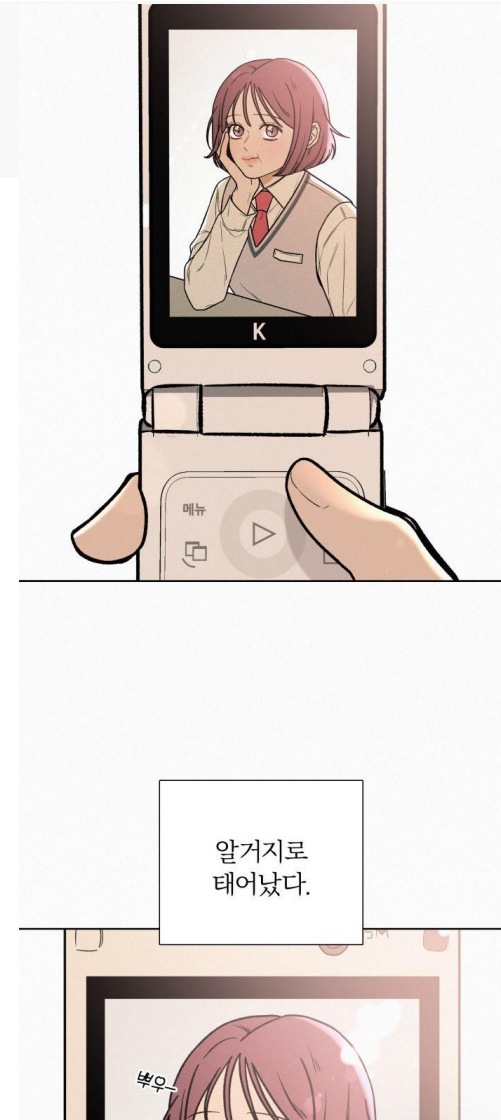
2 데이터 구축- 웹툰

데이터 크롤링

- Try1
로맨스/개그/스릴러/스포츠 등의 다양한
장르 웹툰 50 여개 별 1-10화 크롤링
- > Face detection과 이후에 **실사화를 위한
생성 모델 학습에 어려움을 겪음**



- Try2
**얼굴 그림체가 실제 사람과 유사한 웹툰(로
맨스) 위주로 웹툰 당 1-30화 크롤링**



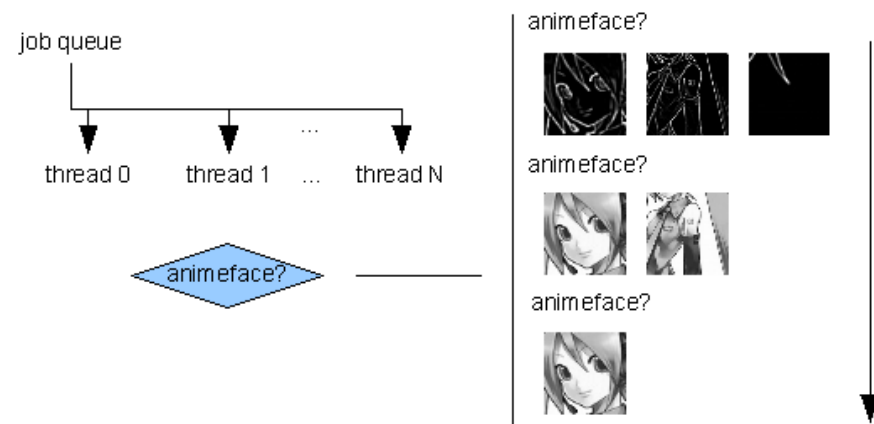
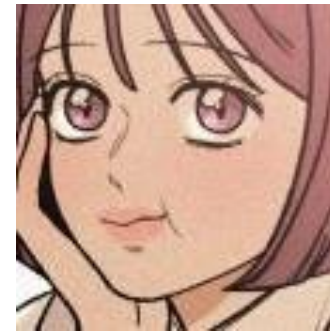
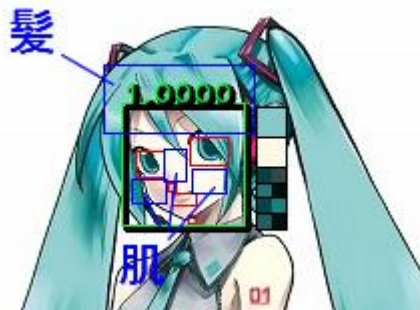
2 데이터 구축- 웹툰

Face Detection

Insightface, lbpcascade,
animeface 등의 모델 시도



캐릭터 얼굴 검출에 가장 적합한
animeface 파이썬 패키지 사용

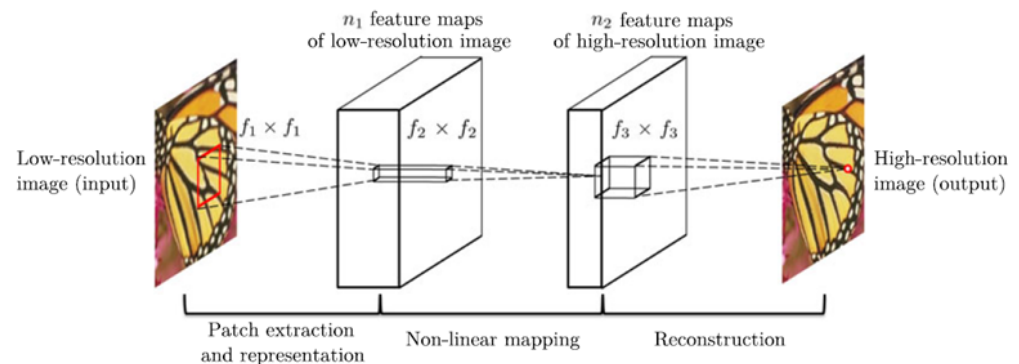
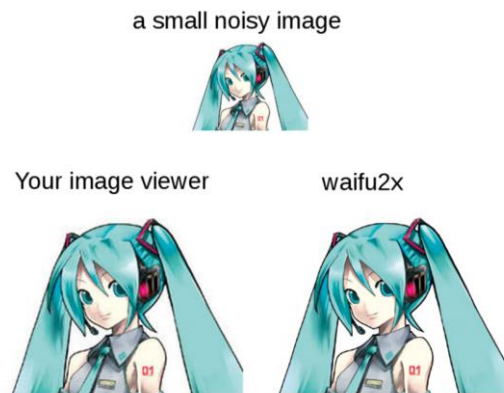
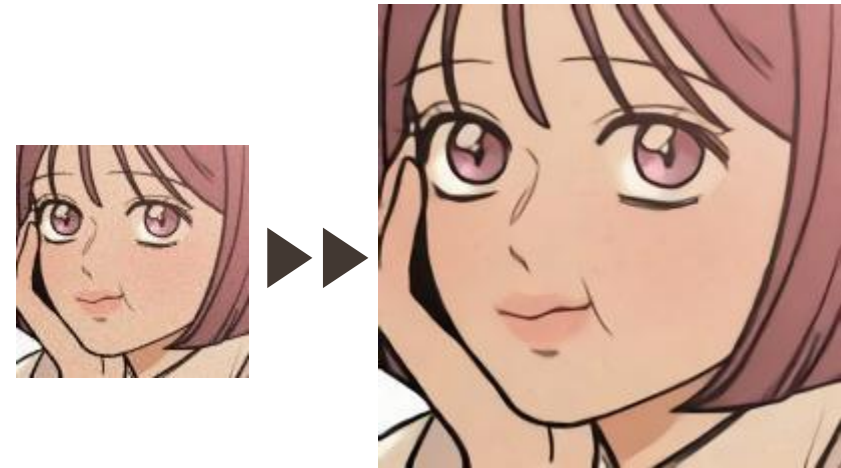


2 데이터 구축- 웹툰

이미지 해상도 높이기

생성 모델의 성능 향상을 위한 이
미지 해상도 높이기 진행

-> SRCNN기반의 **waifu2x** 사용

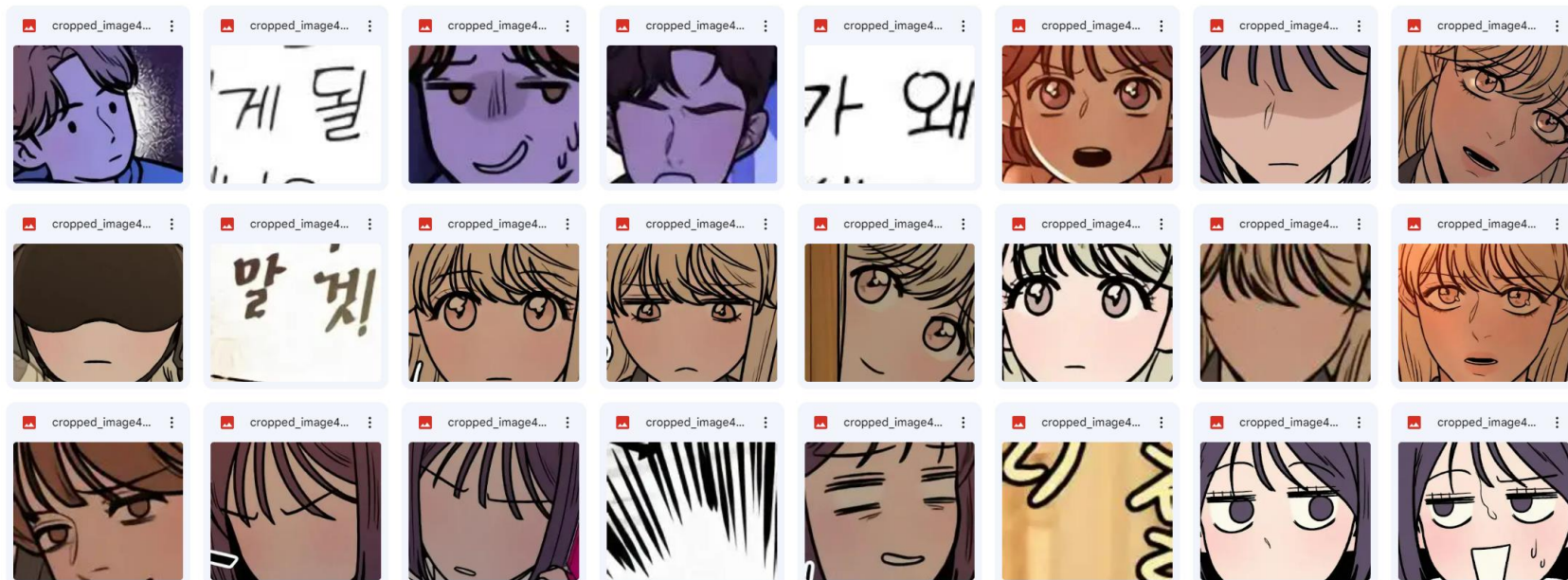


2 데이터 구축- 웹툰

주요 캐릭터 이미지 선별

엑스트라(일회성 등장) <<< 주인공 ▶ 효과적인 feature 학습 유도

-> 가비지 데이터 일부 제거 가능



2 데이터 구축- 웹툰

주요 캐릭터 이미지 선별

feature vector 추출 -> 주인공 이미지와 유사도 비교
-> 유사도 상위 250개 이미지만 저장 -> 수작업으로 제거

- feature vector 추출모델: VGG16, RESNET18 시도
- 유사도: 코사인 유사도, 유클리디안 거리 시도
- > VGG16 X 코사인 유사도



Main char



상위 30개 plot

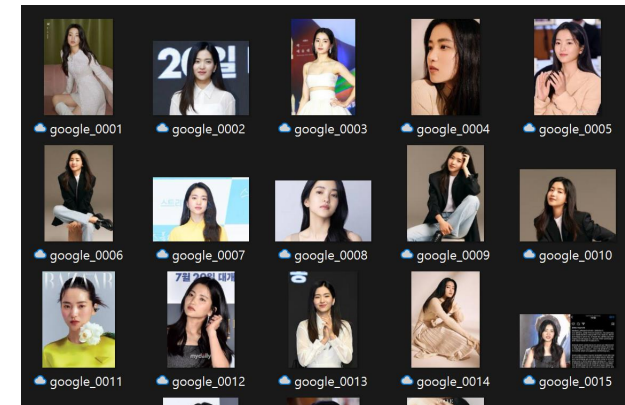
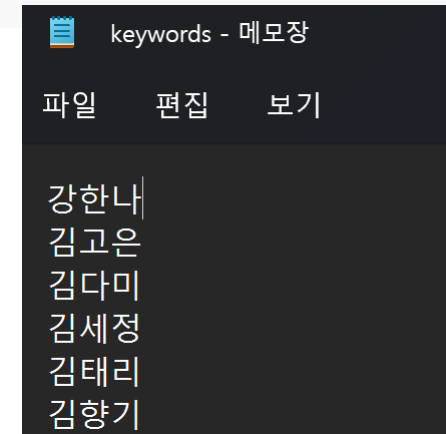
2 데이터 구축- 연예인

데이터 크롤링

- **Auto Crawler** 코드 차용 :
Google, Naver에서 키워드 단위로 이미지를
빠르게 크롤링해주는 멀티프로세싱 코드
- > 동명이인, 관련 인물 등 가비지 데이터가 함께
크롤링되어 1차적으로 제거



- 남자, 여자 연예인 25~30명에 대하여
인당 약 100장의 이미지 크롤링



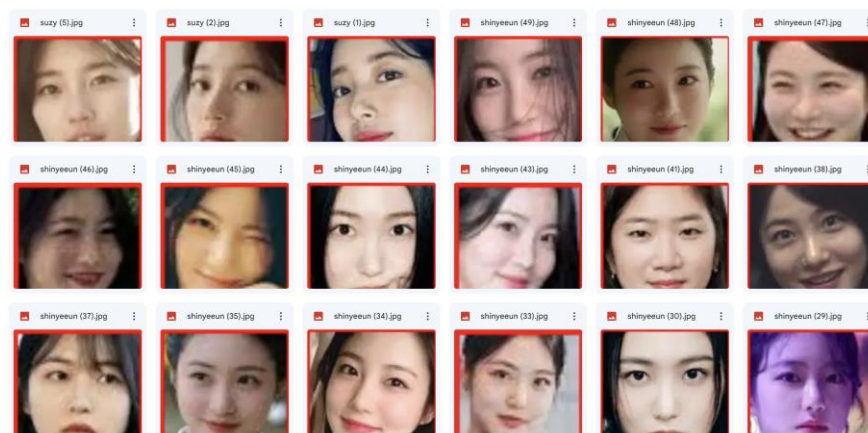
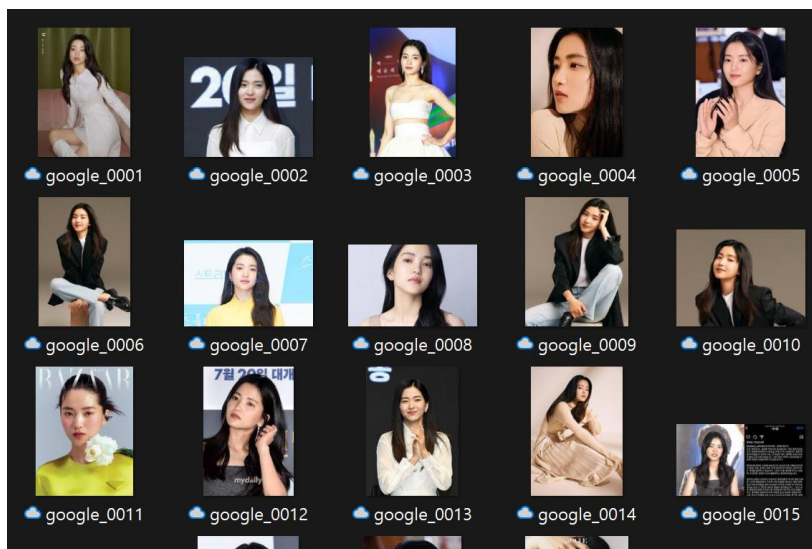
2 데이터 구축- 연예인

얼굴 Crop

Haar Cascade Face Crop 사용:

DSFD, ResNet Face Detector과 같은 NN 모델보다 효율이 높은 ML 기반의 모델

-> 얼굴이 아닌 다른 Artifact가 Crop된 가비지
2차적으로 제거



2 데이터 구축- 연예인

이미지 해상도 높이기

얼굴 Crop 이미지 사이즈 키울 때 해상도 낮아지는 이슈 발생

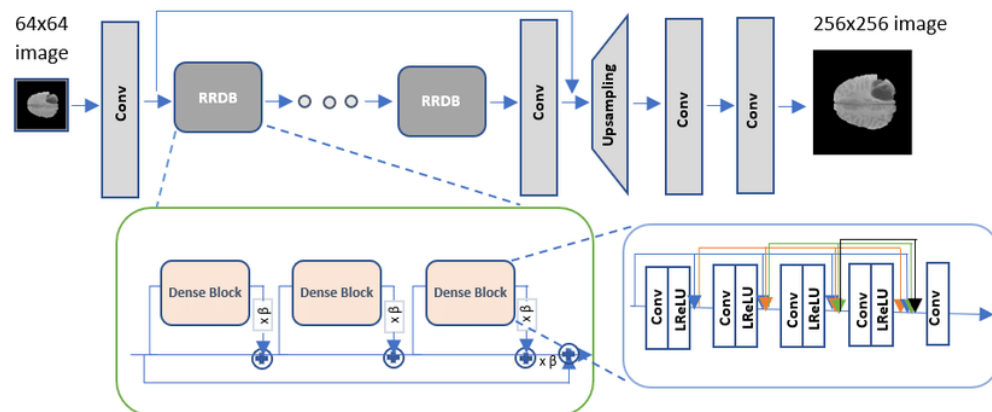


Image Super Resolution,
FSRNet 등의 모델 시도

-> 가장 직관적인 성능이 높은
Real-ESRGAN 사용

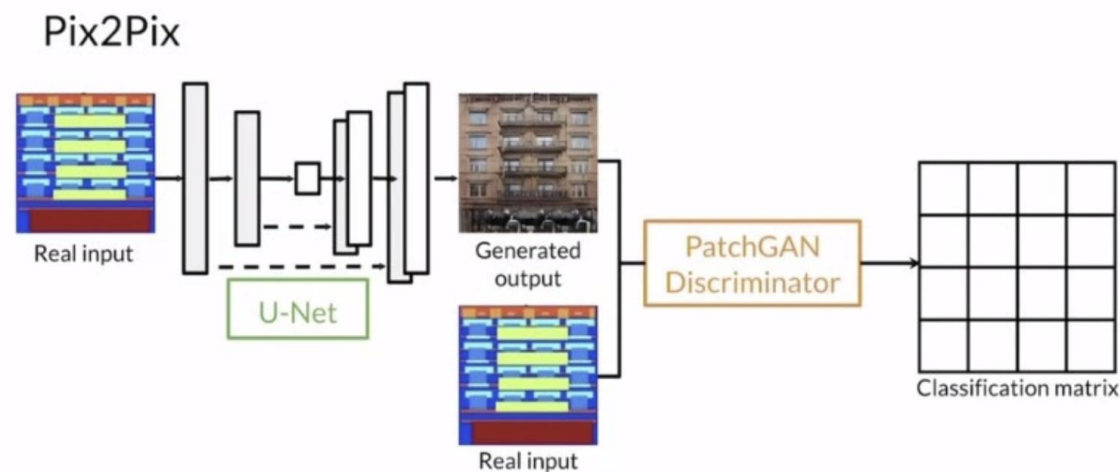


화질을 높여 **Feature Extraction**에
용이하도록 조정



3 모델링 : Unpaired Image2Image Translation

Pix2Pix / CycleGAN / UVCGAN2 / U-GAT-IT



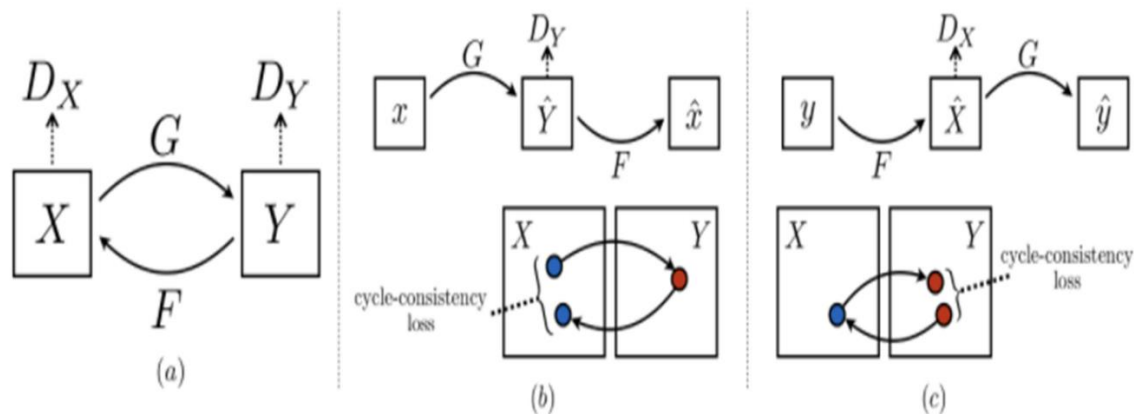
$$\mathcal{L}_{cGAN}(G, D) = \mathbb{E}_{x,y} [\log D(x, y)] + \mathbb{E}_{x,z} [\log(1 - D(x, G(x, z)))],$$

CGAN (Conditional GAN):
- noise를 dropout form의 형태로,
여러층에서 제공될 수 있도록 조정



3 모델링 : Unpaired Image2Image Translation

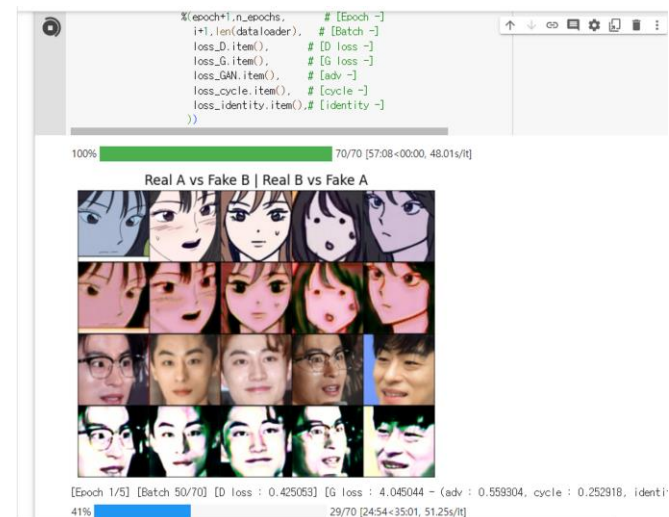
Pix2Pix / CycleGAN / UVCGAN2 / U-GAT-IT



$$\mathcal{L}_{\text{cyc}}(G, F) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\|F(G(x)) - x\|_1] + \mathbb{E}_{y \sim p_{\text{data}}(y)} [\|G(F(y)) - y\|_1].$$

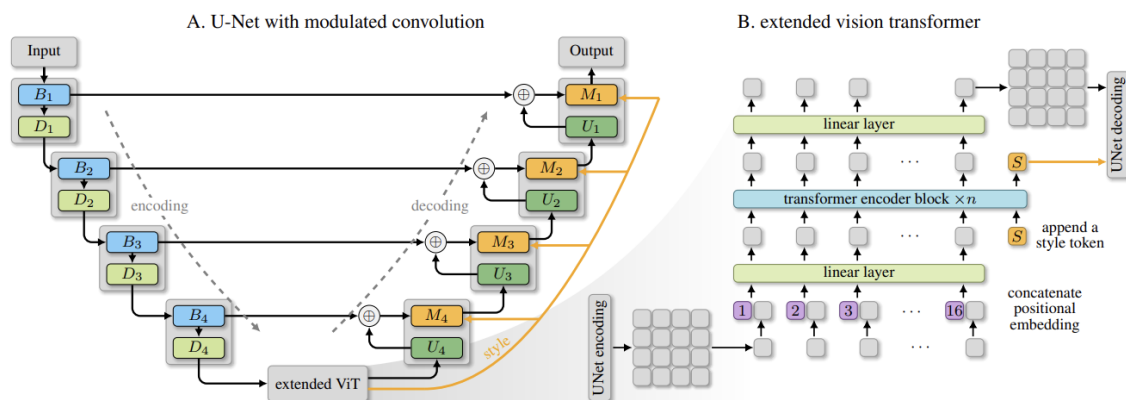
Cycle consistency loss :

- Forward cycle-consistency loss : $x \rightarrow G(x) \rightarrow F(G(x)) \approx x$
- Backward cycle-consistency loss : $y \rightarrow F(y) \rightarrow G(F(y)) \approx y$

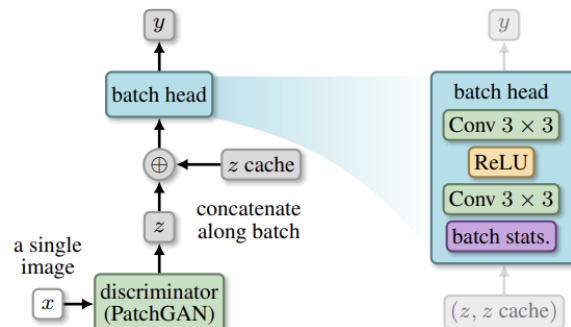


3 모델링 : Unpaired Image2Image Translation

Pix2Pix / CycleGAN / **UVCGAN2** / U-GAT-IT



VS



UVCGAN + eViT

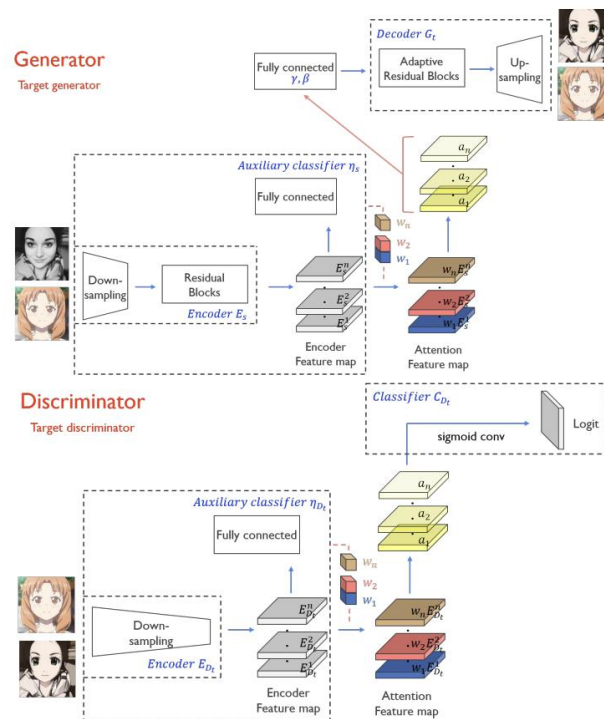
(extended vision transformer):
기존 U-Net 형태의 encoder에
Style Token을 Append하여
decoder로 전달



Iterations : 800+200, - Batch size : 16

3 모델링 : Unpaired Image2Image Translation

Pix2Pix / CycleGAN / UVCGAN2 / U-GAT-IT



- Attention module
- AdaLIN function

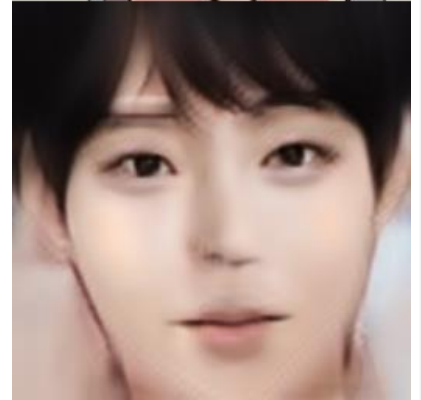
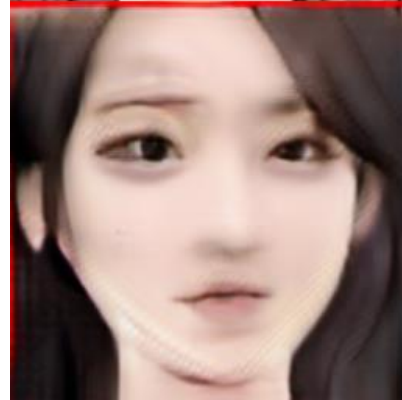
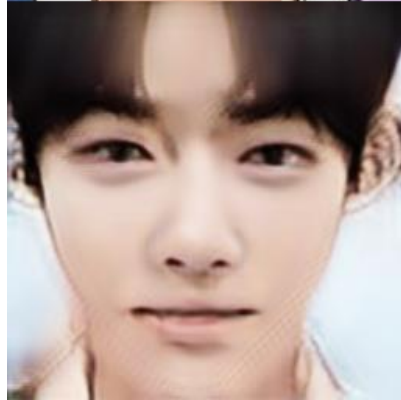
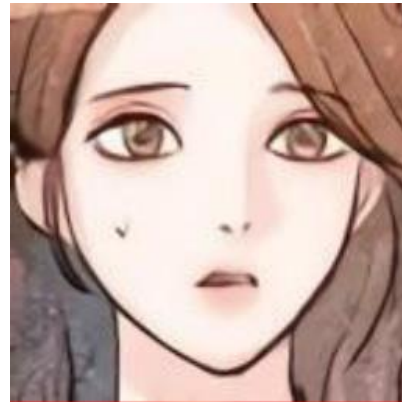
(Adaptive Layer-Instance Normalization) :
Instance Normalization + Layer
Normalization 간의 비율을 Adaptive
하게 선택



3 모델링 : 최종 모델 U-GAT-IT

U-GAT-IT 모델 사용하여 웹툰 → 실물 사진으로 전환 시도

- Iterations : 18000, - Batch size : 2, - Discriminator Layer : 4



3 모델링 : 최종 모델 U-GAT-IT

Input:
Face Detection에 사용된
웹툰 별 주인공 사진



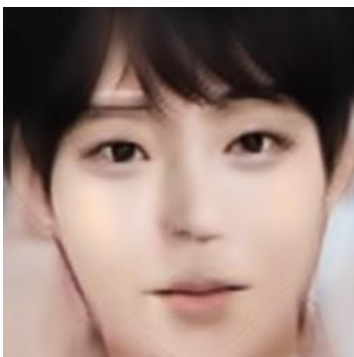
Main char
정면 위주

유사도 비교 대상:
연예인당 정면 위주의 고
화질 사진 2-3장



비교 연예인
정면 위주

U-GAT-IT 사용하여
웹툰 인물 실사화



Output:
Resnet 18 X 코사인 유사도
-> 유사도 top1 연예인 선정



Webtoon Image
web_men1.jpg



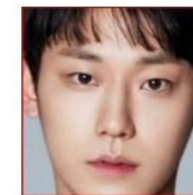
Similarity value 0.891
men_celeb (56).jpg



Similarity value 0.875
men_celeb (27).jpg



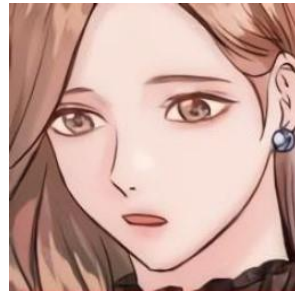
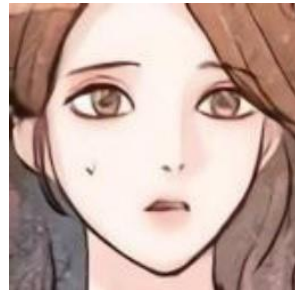
Similarity value 0.865
men_celeb (67).jpg



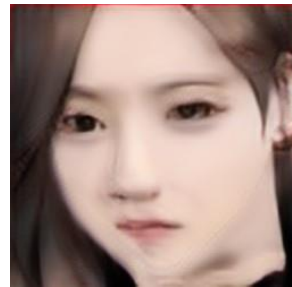
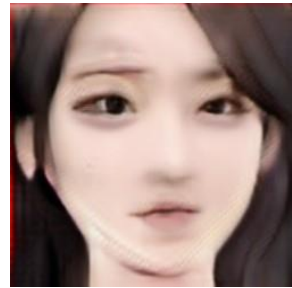
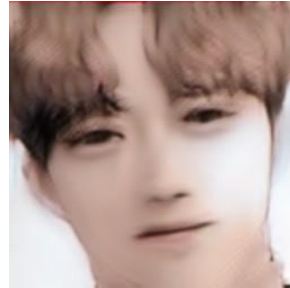
Similarity value 0.865
men_ (6).jpg

3 모델링 : 최종 모델 U-GAT-IT

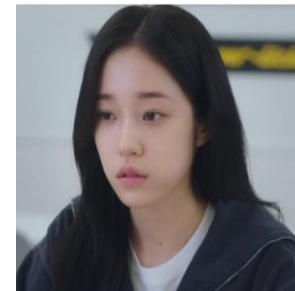
Webtoon



실사화



Celebrity



동일 웹툰
캐릭터



동일 연예인

4 느낀점 및 한계

CV 분야를 전반적으로 공부할 수 있었던 시간이었다.

데이터 구축부터 모델링까지 직접 준비해 가면서

데이터 구축과 팀워크 등의 중요성을 느낄 수 있었다.

컴퓨팅 환경의 한계로 인해 더 다양한 모델링을

시도해 보지 못한 점이 한가지 아쉬운 부분이다.

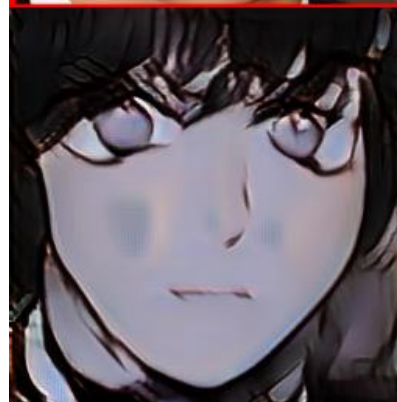
-> 남/여 성별을 구분하여 모델 학습을 진행한 경우의 결과가

더 우수할 것 같아 시도해 보고 싶은 생각이 든다.

* 번외

U-GAT-IT 모델 사용하여 실물 → 웹툰 사진으로 전환 시도

세부 사항 : 전과 동일, - Iterations : 18000, - Batch size : 2, - Discriminator Layer : 4

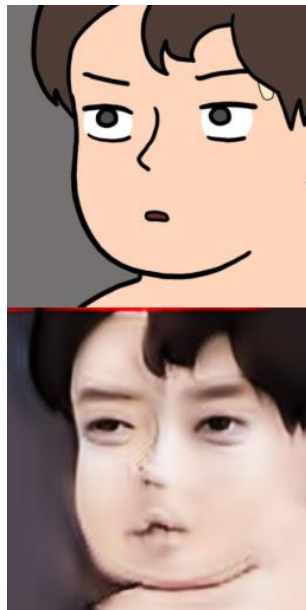


* 번외

개그 웹툰 실사화 시도



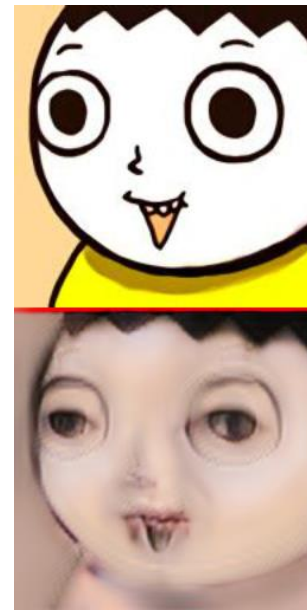
UVCGAN2



U-GAT-IT



UVCGAN2

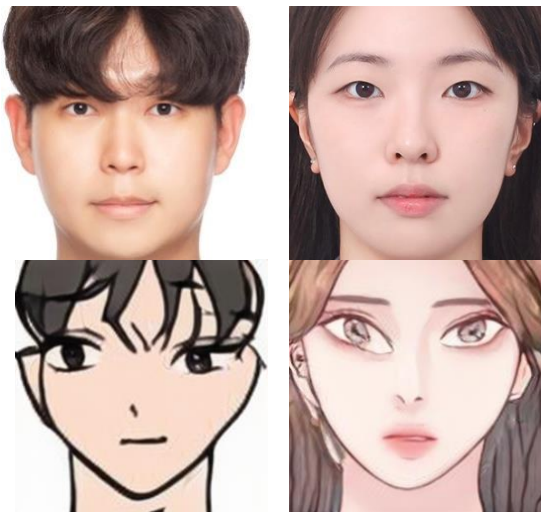


U-GAT-IT

* 번외

우리 팀 웹툰화 시도

UVCGAN2



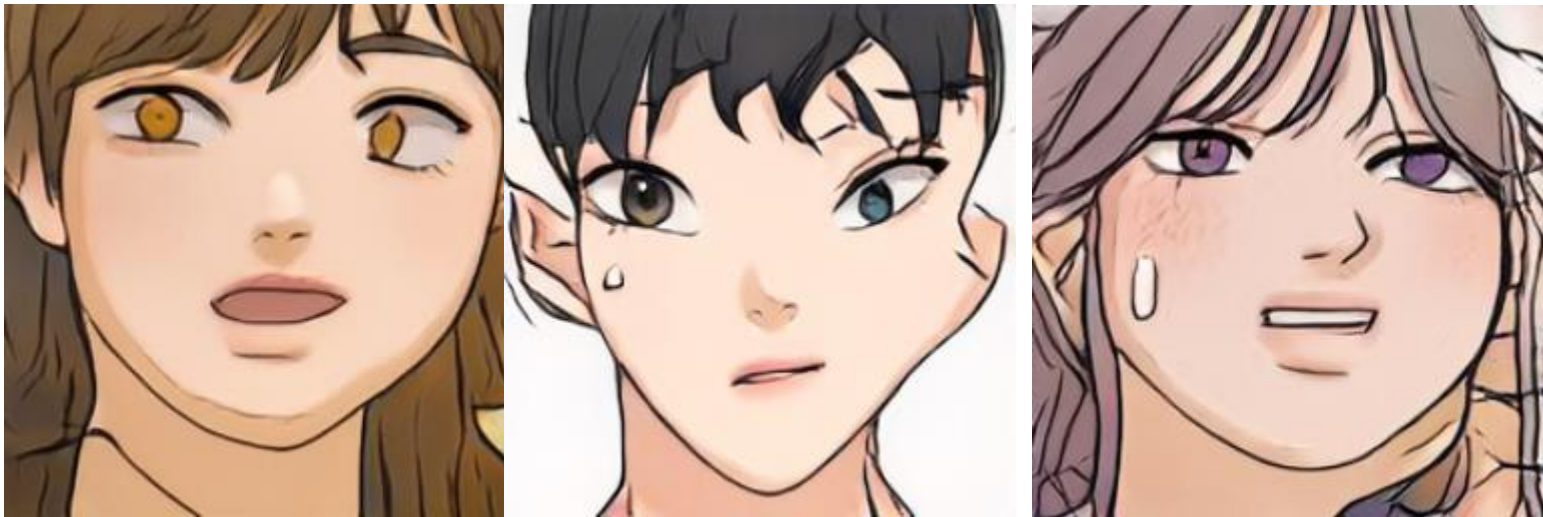
U-GAT-IT



* 번외

쿠빅이들 웹툰화 시도

UVCGAN2



감사합니다.

BANG



POW!