

• • • • •

적대적 생성 신경망 기반 가상 피팅 서비스

Team 무딱대(무취딱) AI융합학부 20201809 정재윤
AI융합학부 20201812 한승우



INDEX

01 #ONE

기획 의도 및 시장 배경

온라인 쇼핑몰의 시장 상황에 대한 설명을 통해, 프로그램의 기획 의도에 대해 소개한다.

02 #TWO

시장과의 차별성

시장에 이미 출시된 상품들과 비교해보며, 본 제품의 차별성을 소개한다.

03 #THREE

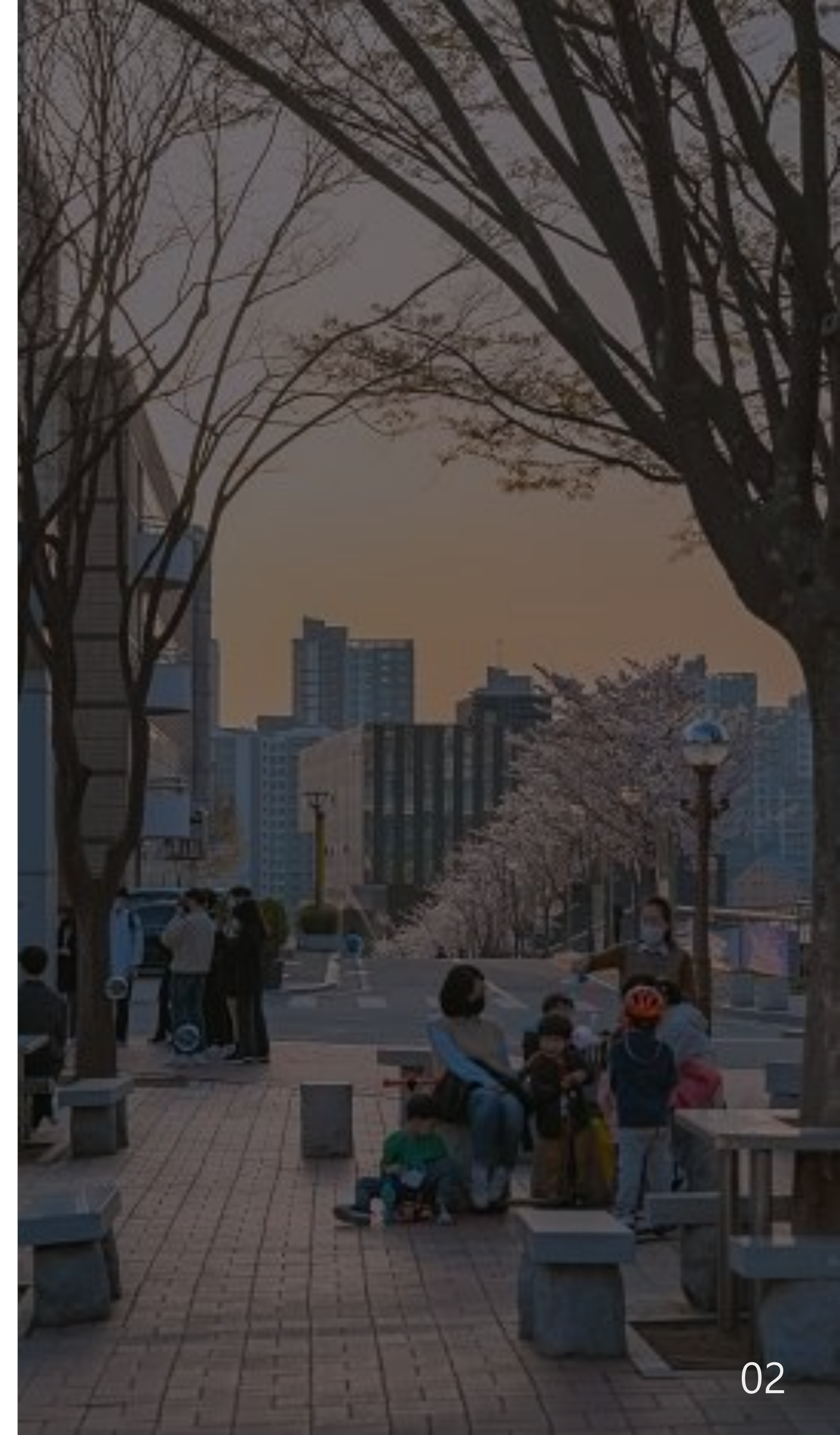
모델 및 파이프라인

Generation 모델로 사용한 HR-VITON 및 Generation 모델을 사용하기 위해 Preprocess 과정에 사용한 모델을 소개한다.

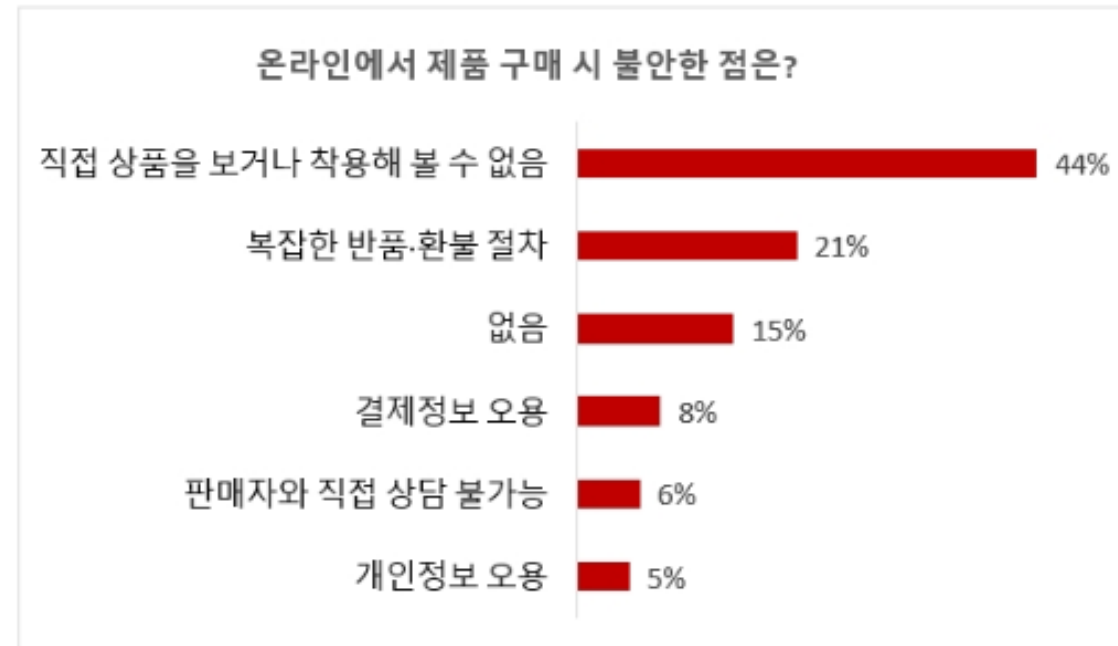
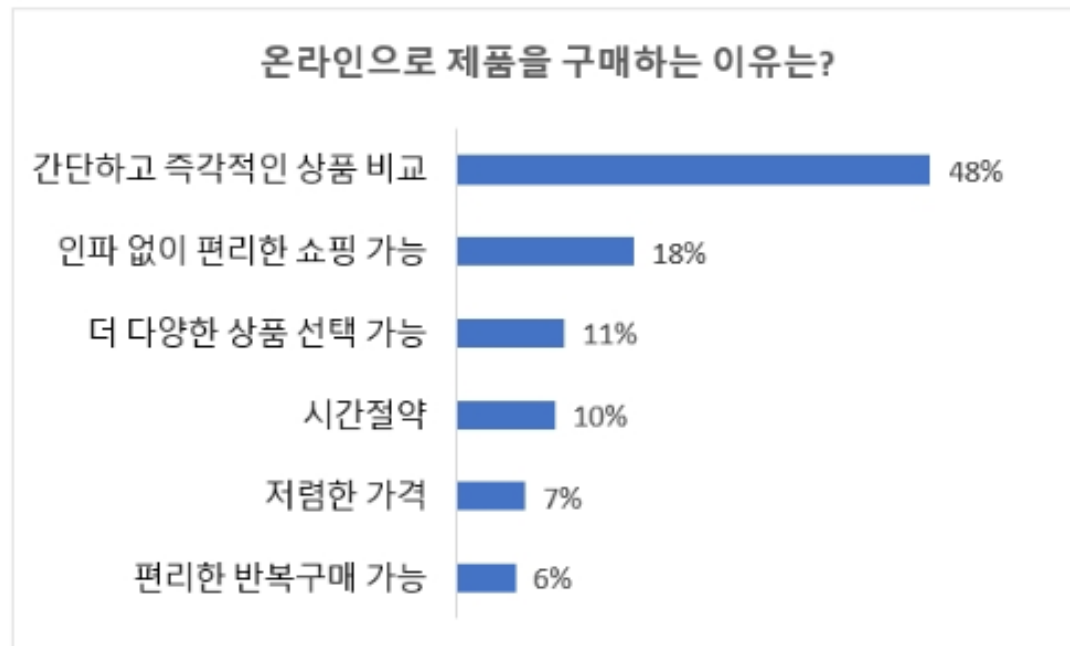
04 #FOUR

기대효과 및 서비스 시연

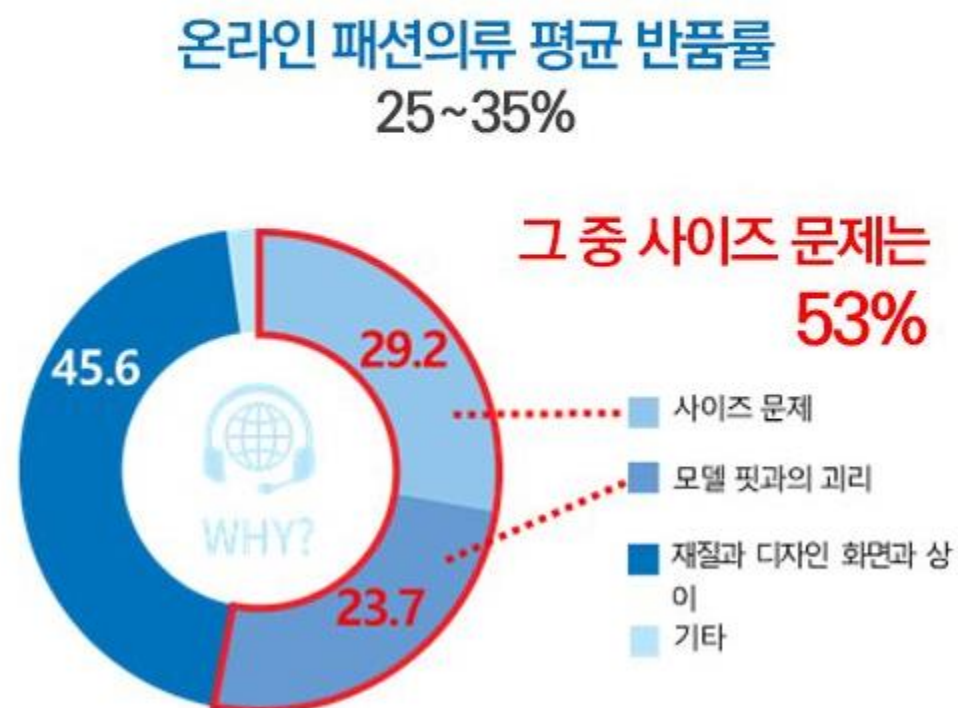
이러한 모델로 온라인 쇼핑몰 기업에서 얻을 수 있는 기대효과 및 사업화 전략을 소개하고, 서비스 시연을 하며 마친다.



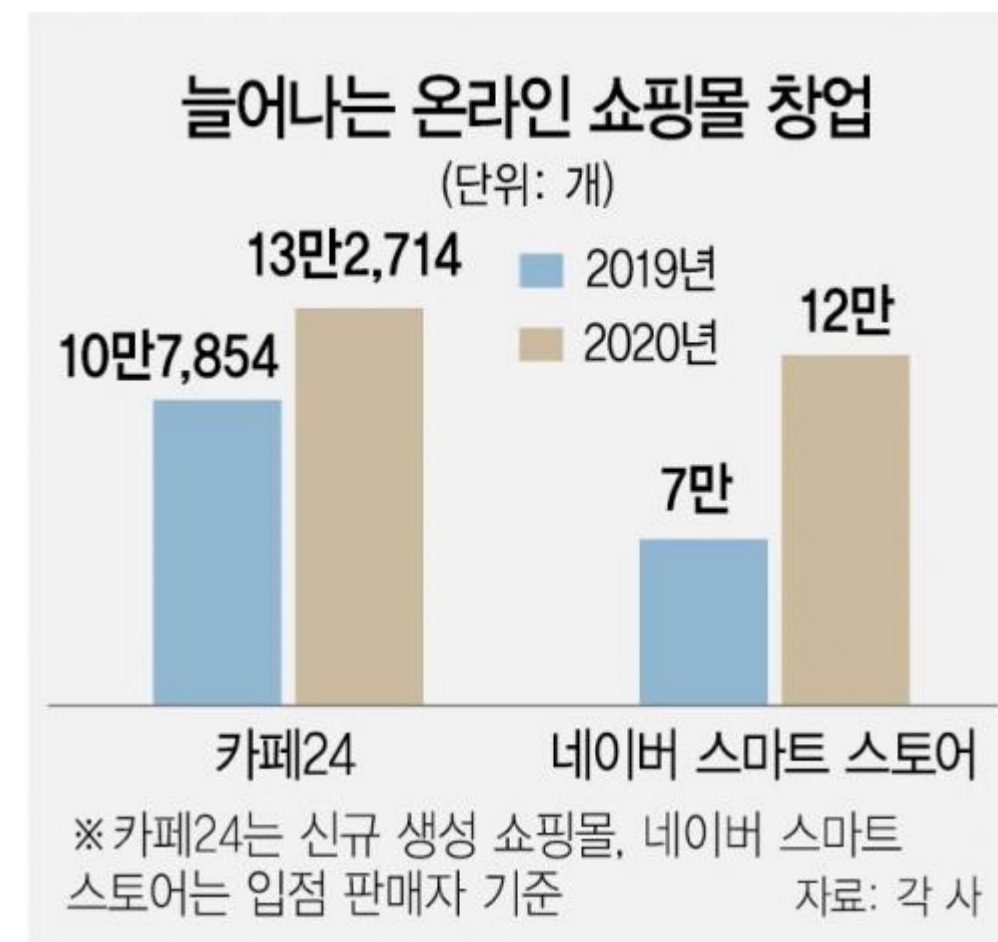
01. 기획 의도 및 시장 배경



국내 온라인 의류쇼핑 시장
14조 2천억원

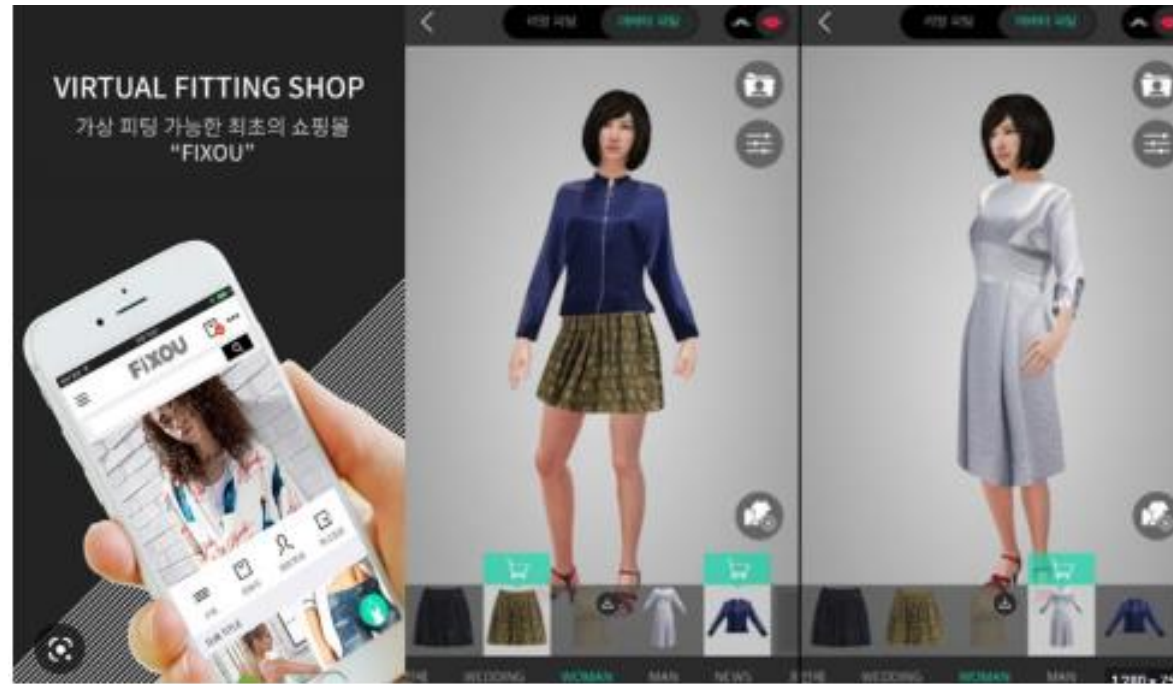


- 온라인 쇼핑은, 간단하게 제품을 구매할 수 있다는 장점이 있지만, 직접 상품을 보거나 착용해 볼 수 없어, 환불 및 반품의 비중이 높다는 단점이 존재함.



- 이러한 단점에도 불구하고, 편리하다는 점 때문에 온라인 쇼핑몰의 파이가 점점 커지고 있다는 것을 통계를 통해 확인할 수 있음.

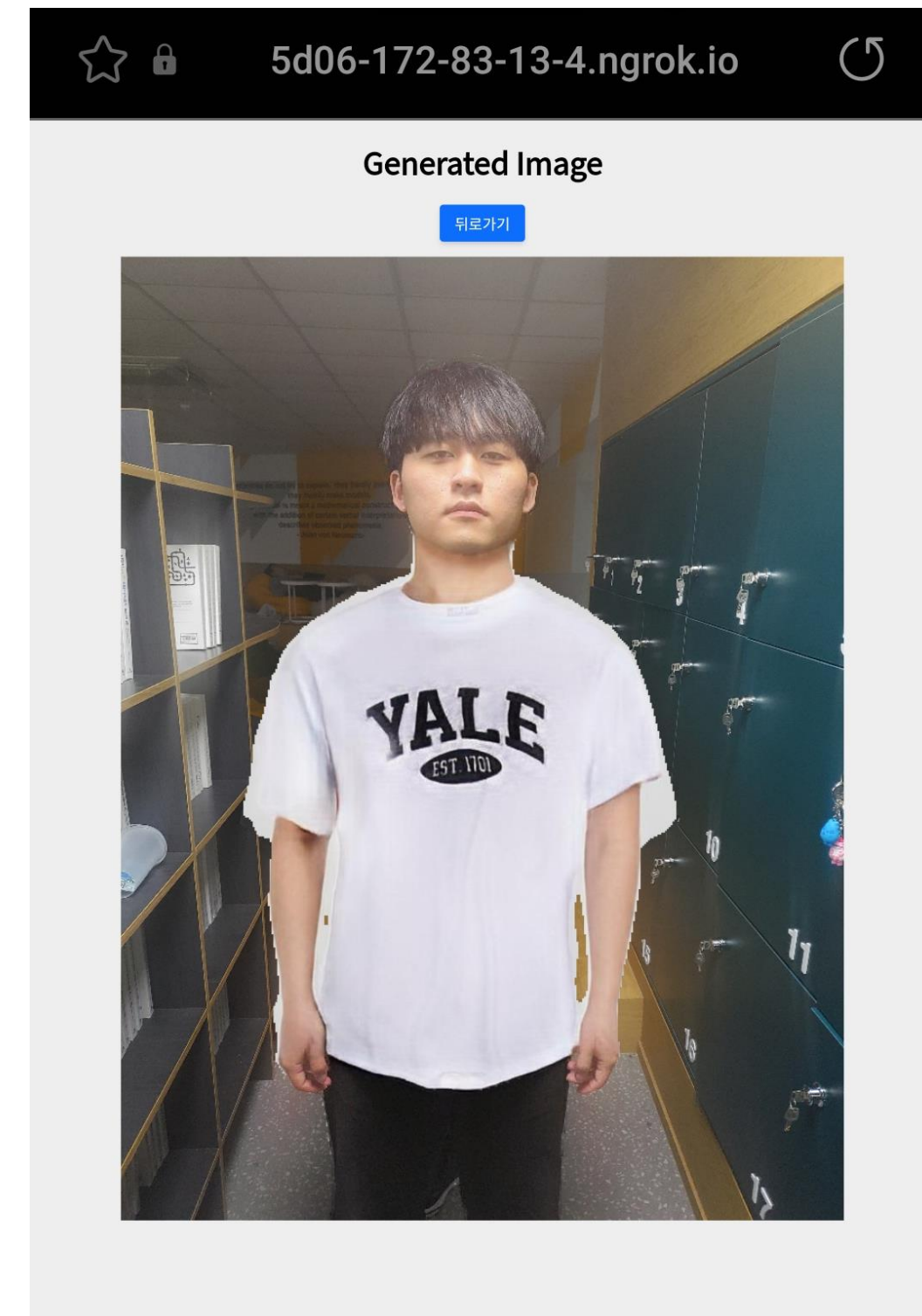
02. 시장과의 차별성



- FIXOU라는 앱은 내가 원하는 옷을 고를 수 있으나, 그 대상이 자신이 아닌 아바타라는 점에서 한계가 존재.



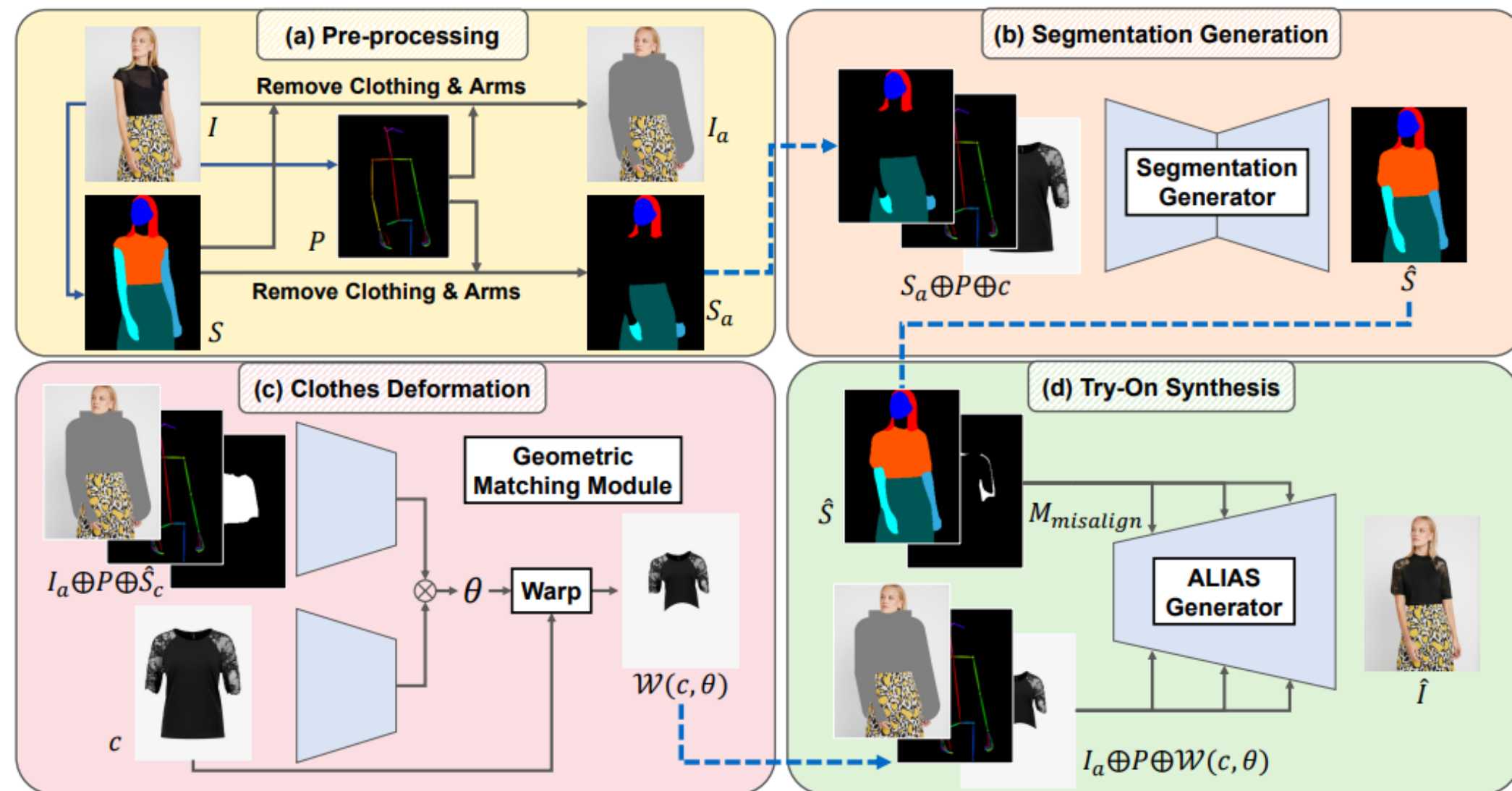
- ETRI에서 개발한 3D 기반 가상 피팅 시스템은, 3D 센서가 필요하기 때문에, 해당 제품이 구비된 매장으로 직접 찾아가야 함.



- WEB을 통해 간편하게 접근할 수 있으며, 아바타가 아닌 나 자신에게 옷을 입힐수 있다는 장점이 모두 존재함.

02. 모델 및 파이프라인

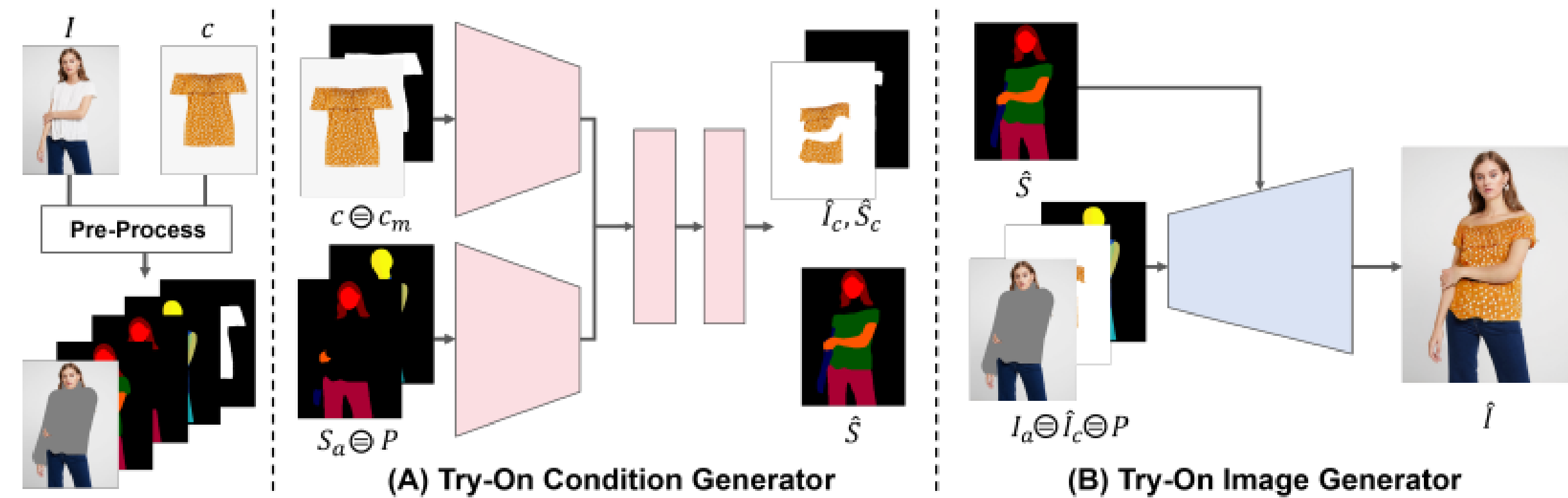
Base Model : HR-VITON



VITON-HD: High-Resolution Virtual Try-On via Misalignment-Aware Normalization, Sep 2021

Segmentation generation과 class deformation model을 각각 분리하여 이미지를 생성하는 모델

오른쪽 사진은, 모델을 분리함으로 인해 발생한 misalignment이다.

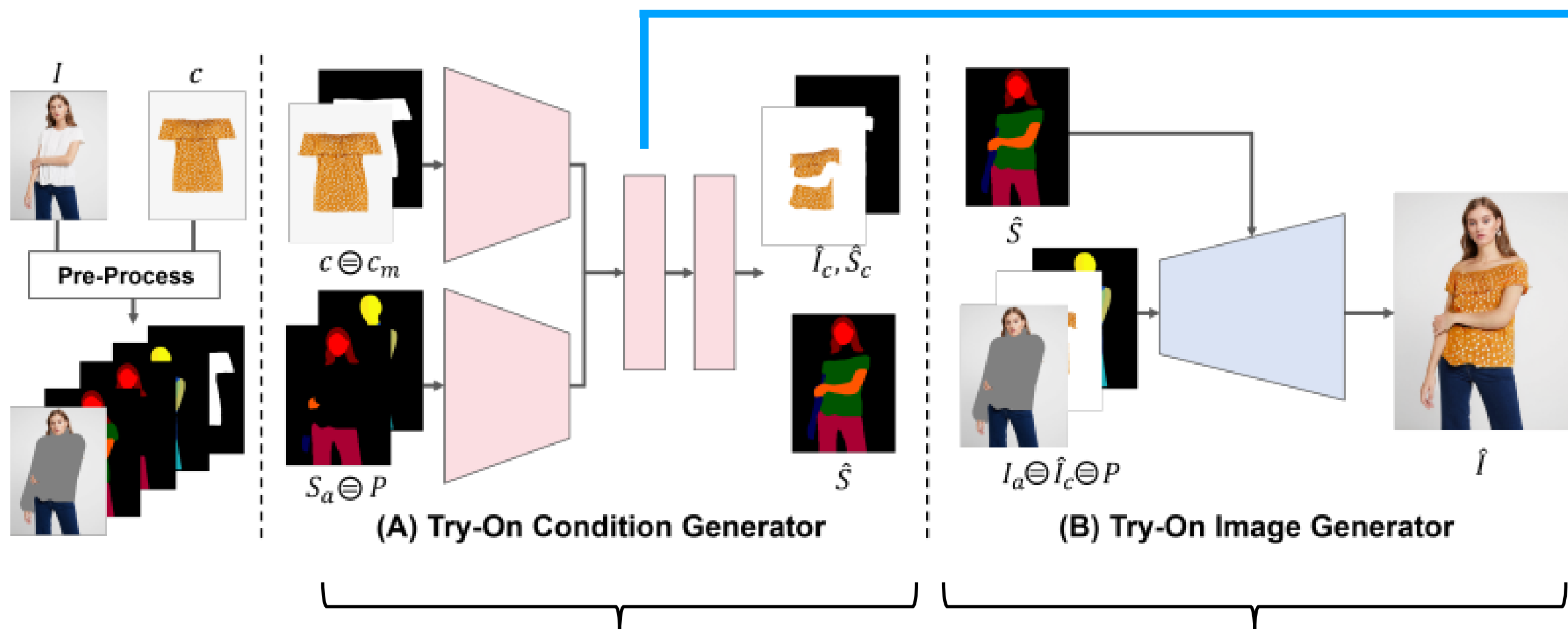


HR-VITON : High-Resolution Virtual Try-On with Misalignment and Occlusion-Handled Conditions. Jul 2022

Image 및 segmentation을 하나의 모델로 통합하여, train, inference 속도 및 accuracy를 높임

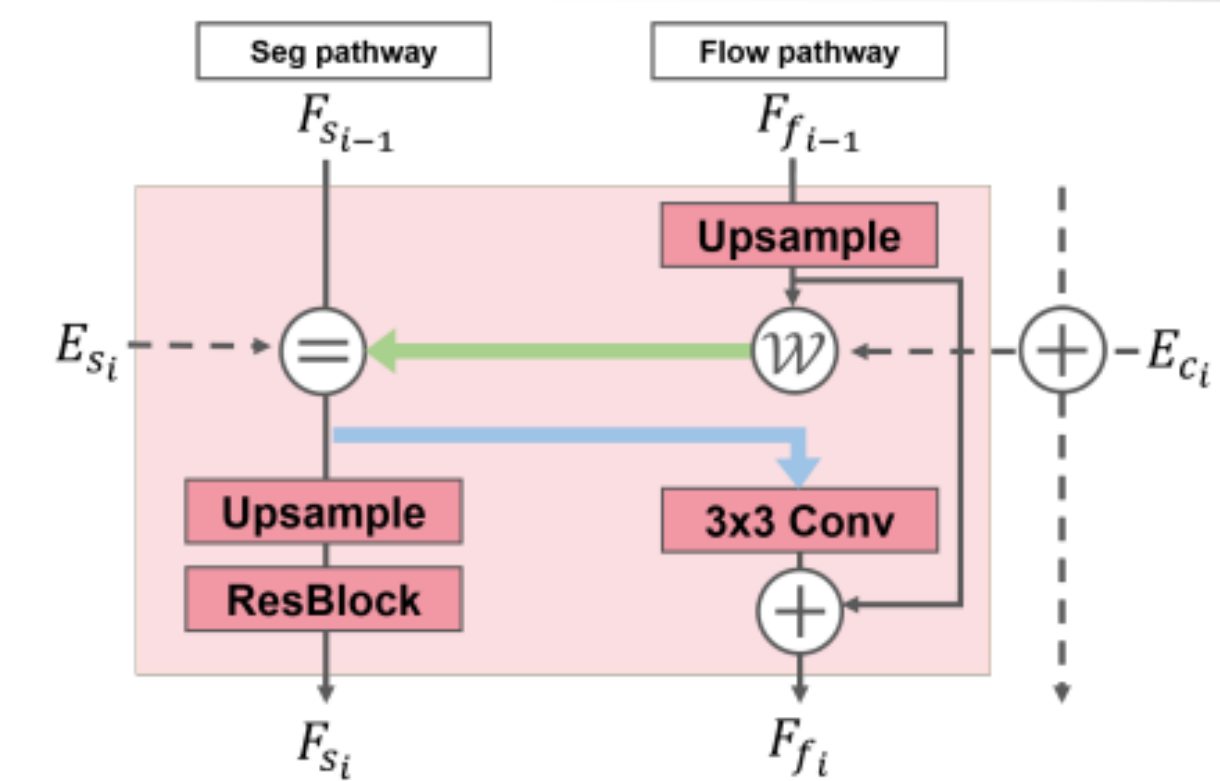
02. 모델 및 파이프라인

Base Model : HR-VITON



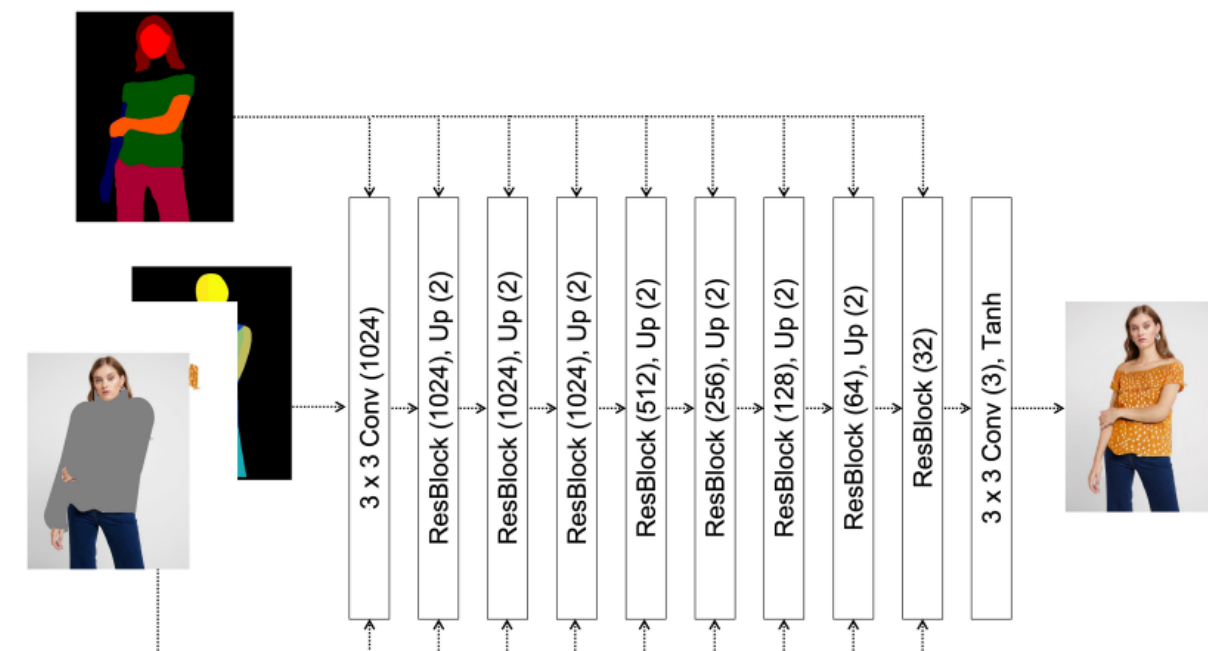
옷과 사용자의 사진을 fusion하는 과정을 통해, 변형된 옷과 segmentation을 generate

변형된 옷과 segmentation을 통해 최종 결과 이미지를 generate



Feature fusion block

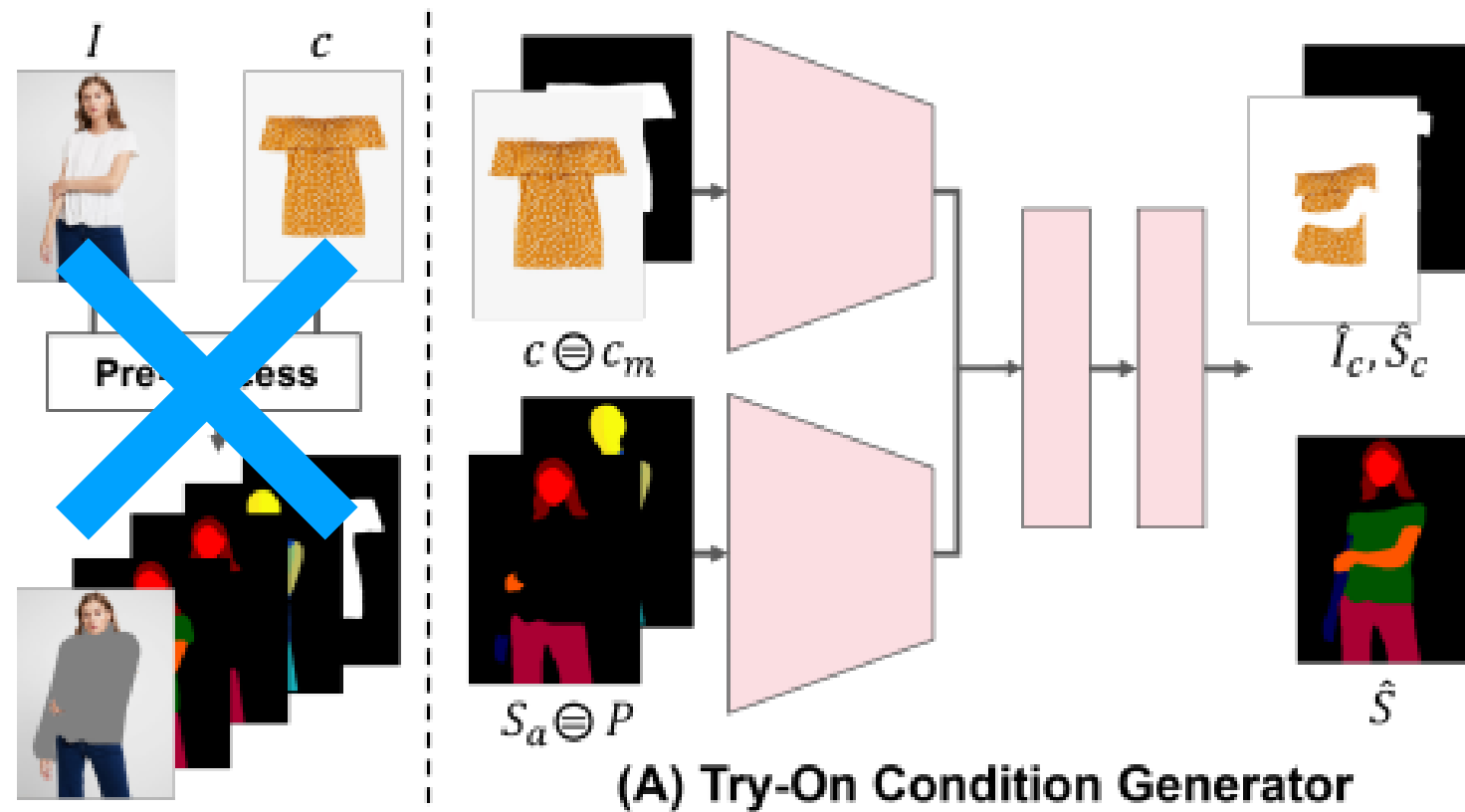
-> concatenate 및 add 연산을 통해, 옷에 해당하는 필터와 segmentation에 해당하는 filter를 섞어줍니다. (기억력이 1인 lstm model과 비슷)



Try-On Image Generator

-> 단순히 upscale ResBlock을 연결한 generator이며, pix2pixHD와 같은 loss function을 통해 학습시킴.

Base Model : HR-VITON (Generation Model)

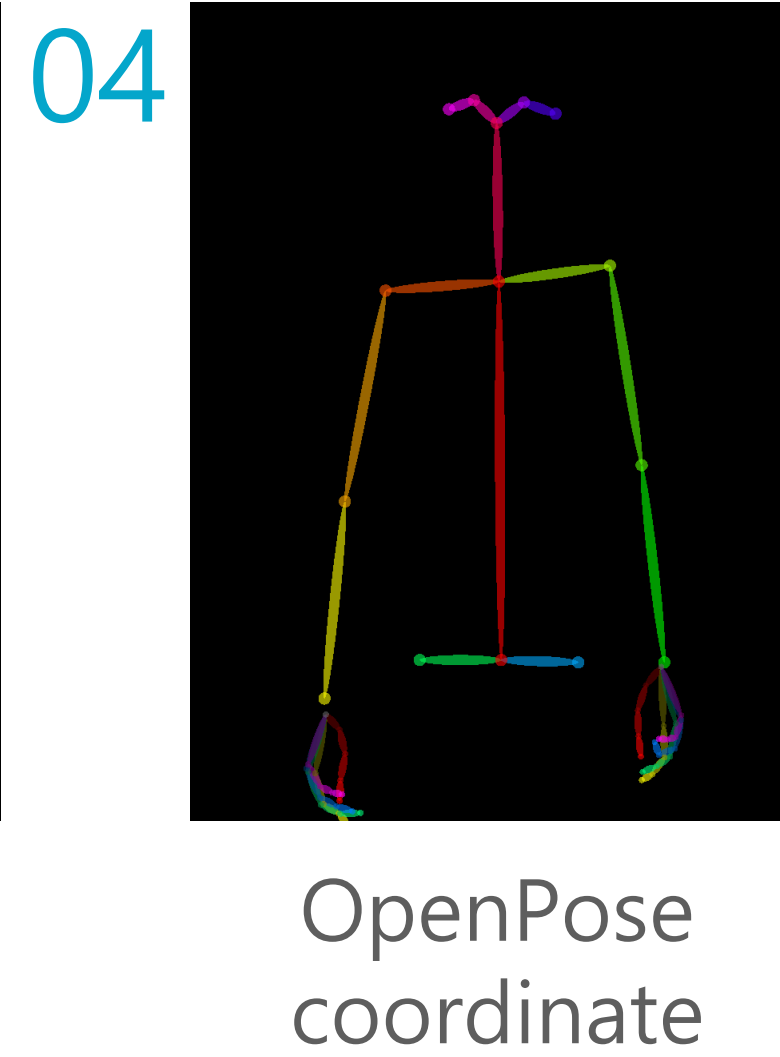
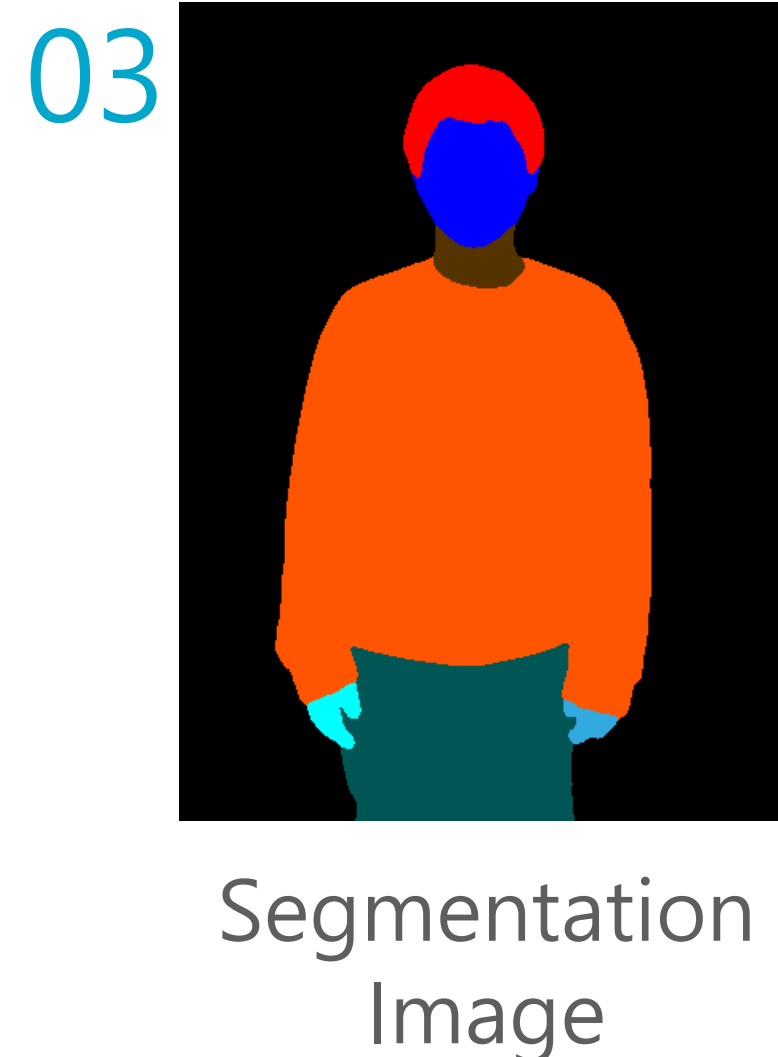
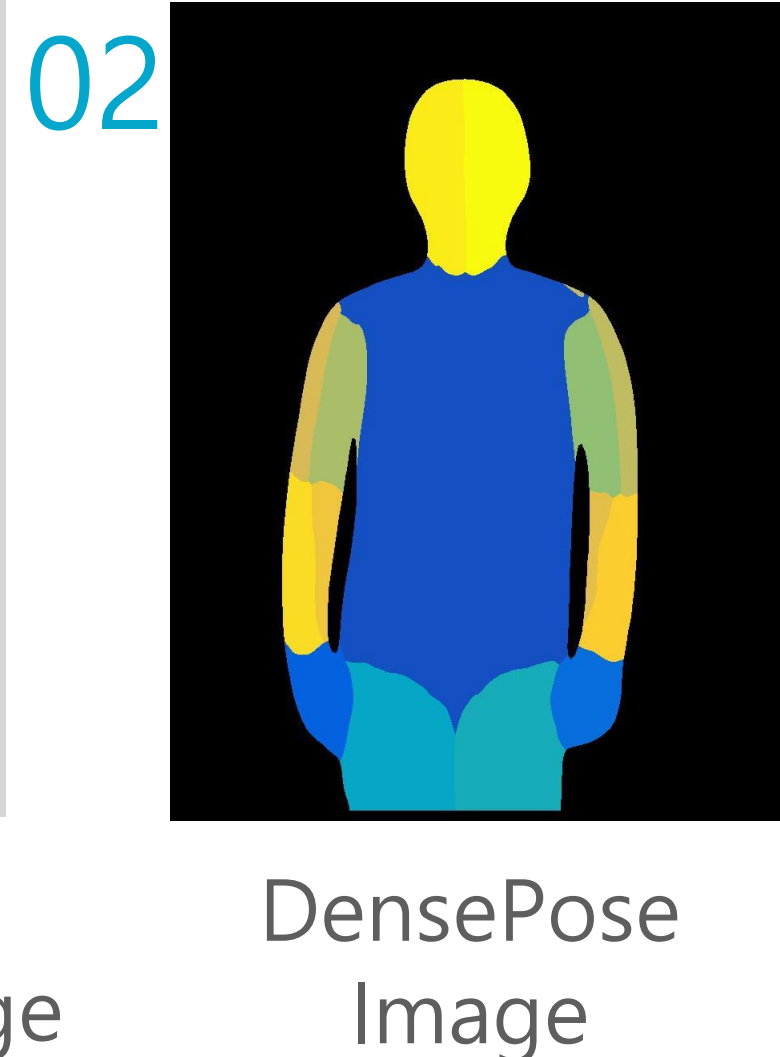
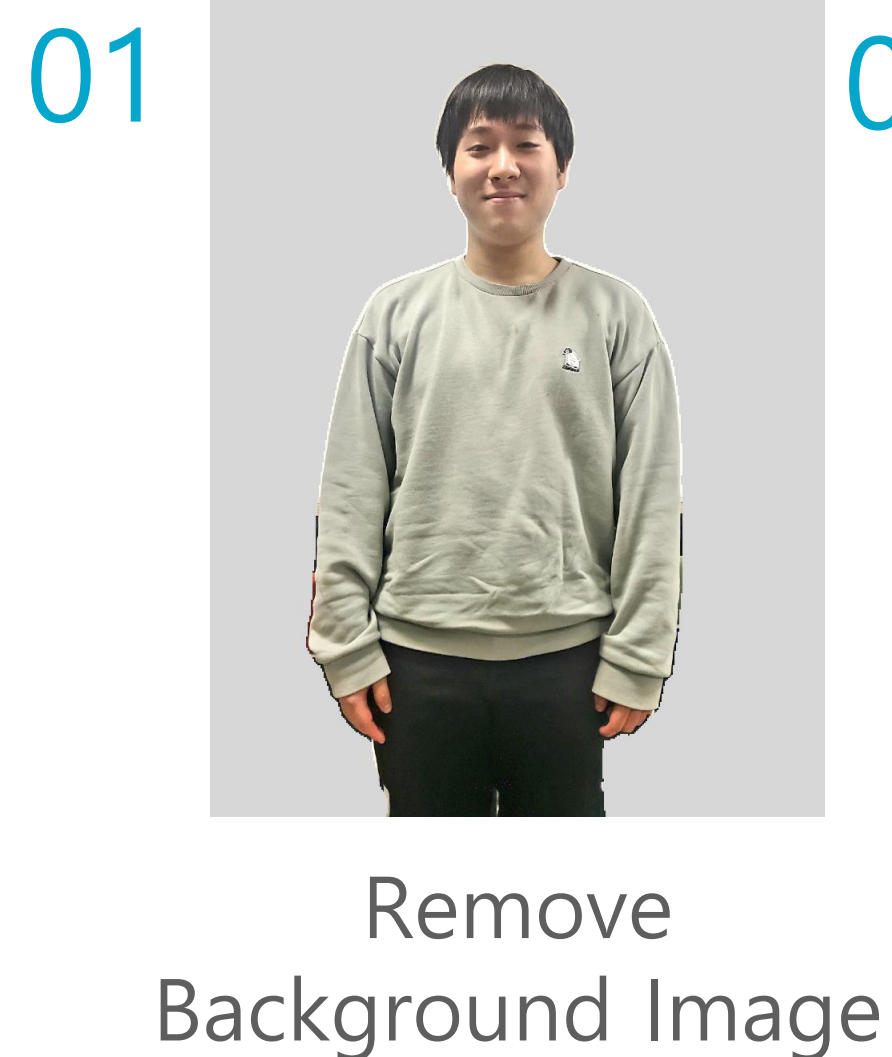
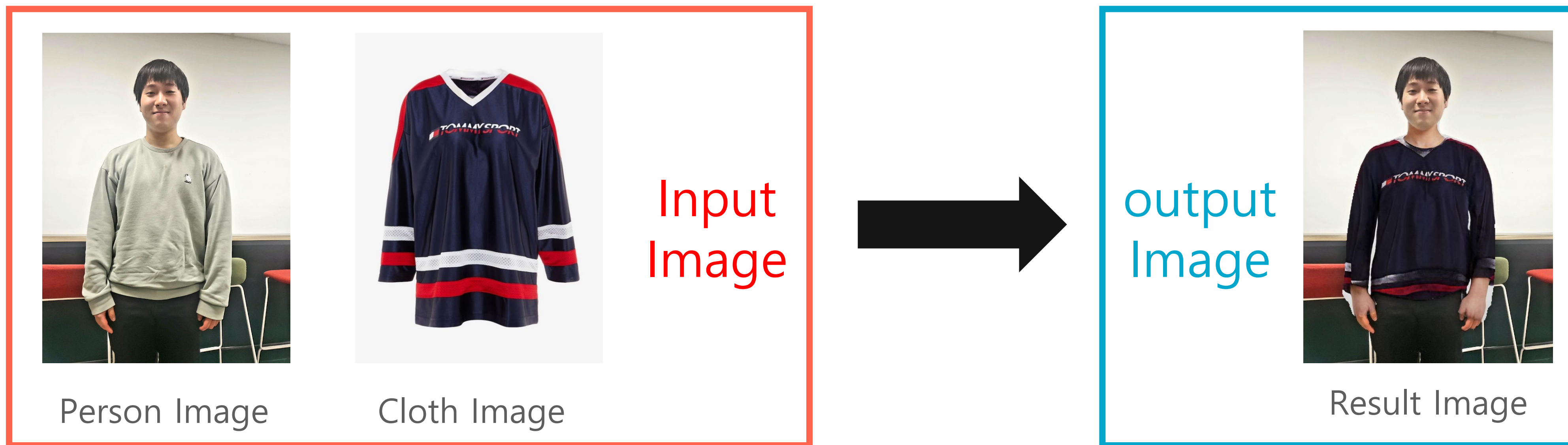


- 그저 라이브러리를 불러와서 pretrained된 모델을 쓰기만 하면 Try-on condition generator와 image generator만 구현되어 있기 때문에 우리가 찍은 사진을 쓰지 못한다.
- 따라서 추가적으로 input으로 쓰이는 사진을 우리가 바꾸고 싶다면 논문 사이사이에 쓰이는 부분과, preprocess 부분 구현이 필요하다.

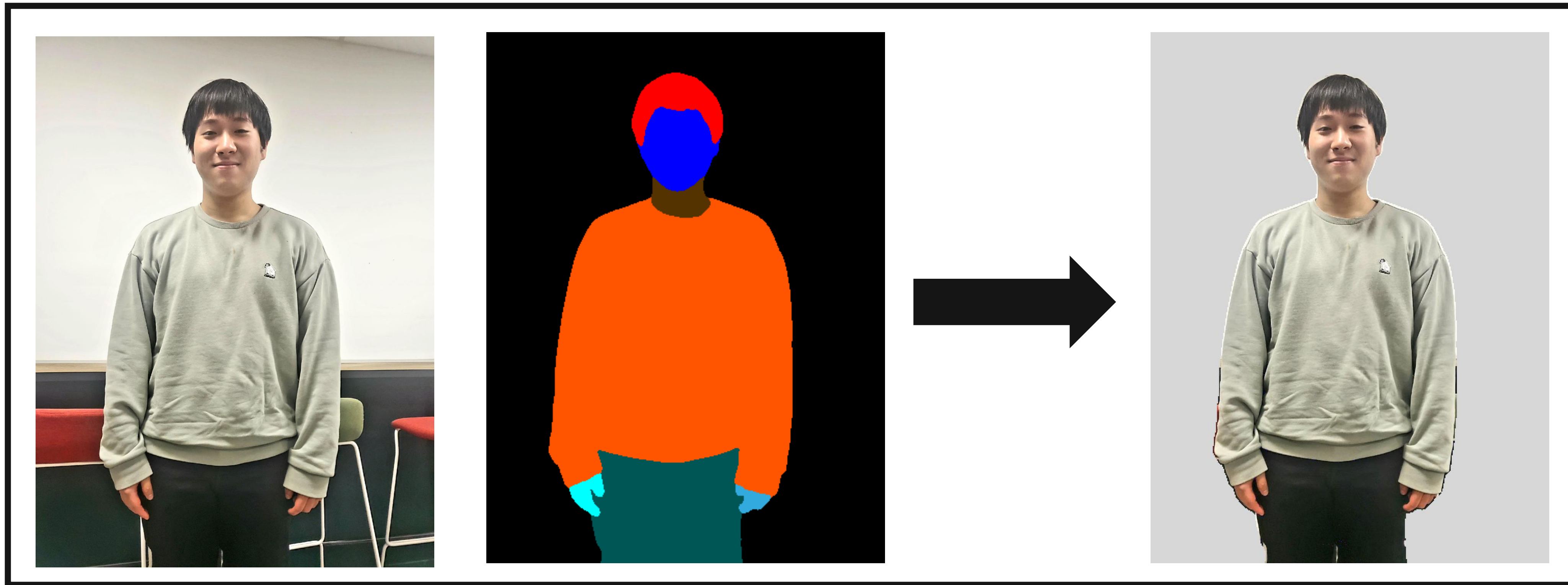
- 최근에도 official 논문 코드구현 github issue 란에 무수한 preprocessing 방법 및 논문결과 재현에 방법에 대한 질문이 올라오고 있다.
- 따라서 HR-VITON의 논문결과를 코드와 모델을 추가하여 재현하고, 우리가 찍은 이미지를 쓸 수도 있도록 데이터 파이프라인을 구축해 실제 TRY-ON 웹사이트를 구축하는 것을 목표로 진행하였다.

- ⊙ Data Preprocessing!
#35 opened 17 hours ago by netuserjun
- ⊙ Offset Generated During Training
#34 opened 8 days ago by Atotem
- ⊙ Help me with transform new images to process
#33 opened 17 days ago by julianorinaldi
- ⊙ How to test with news persons and clothes?
#32 opened 23 days ago by julianorinaldi
- ⊙ draw openpose
#31 opened 28 days ago by tiamo405
- ⊙ How to test HR-Viton on 2 images.
#30 opened last month by ahmadmustafaanis

Images that need to be ready for the base model



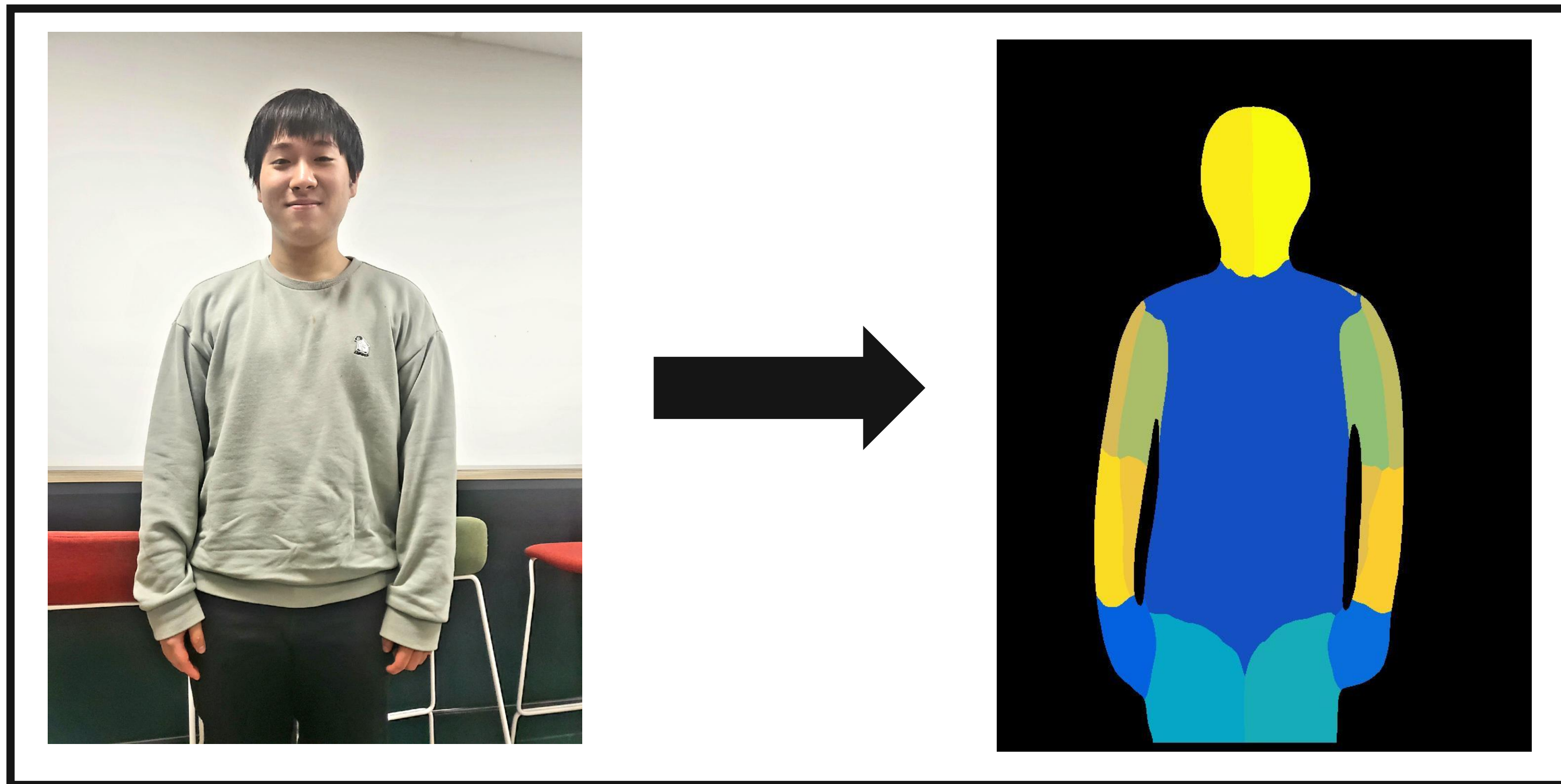
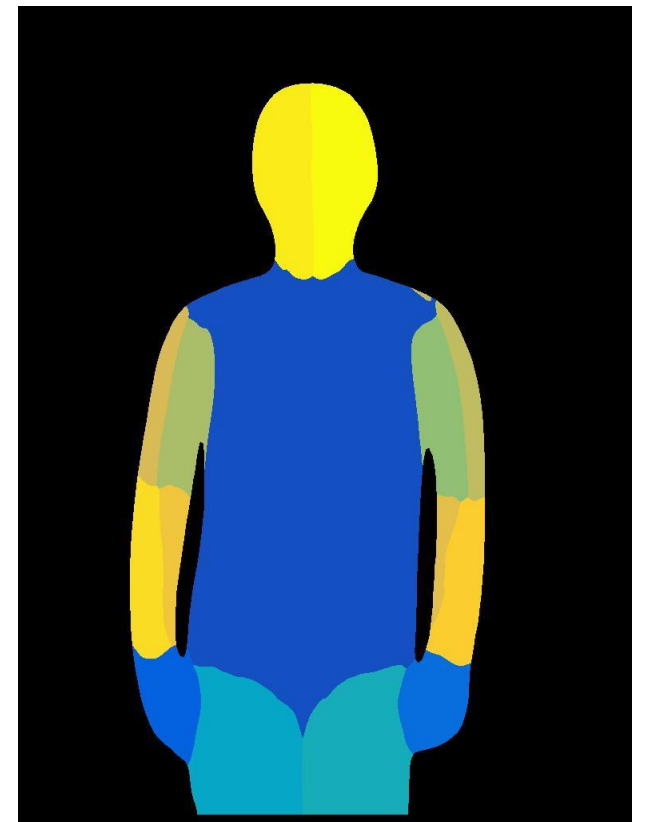
Preprocess – Image 01, Remove Background Image 01



- Segmentation된 mask Image를 이용해서 백그라운드를 제거하고, 백그라운드를 VITON-HD 이미지 셋과 유사하도록 회색으로 채움

Preprocess – Image 02, DensePose Image

02



koo616 commented [on Aug 8](#)

Collaborator

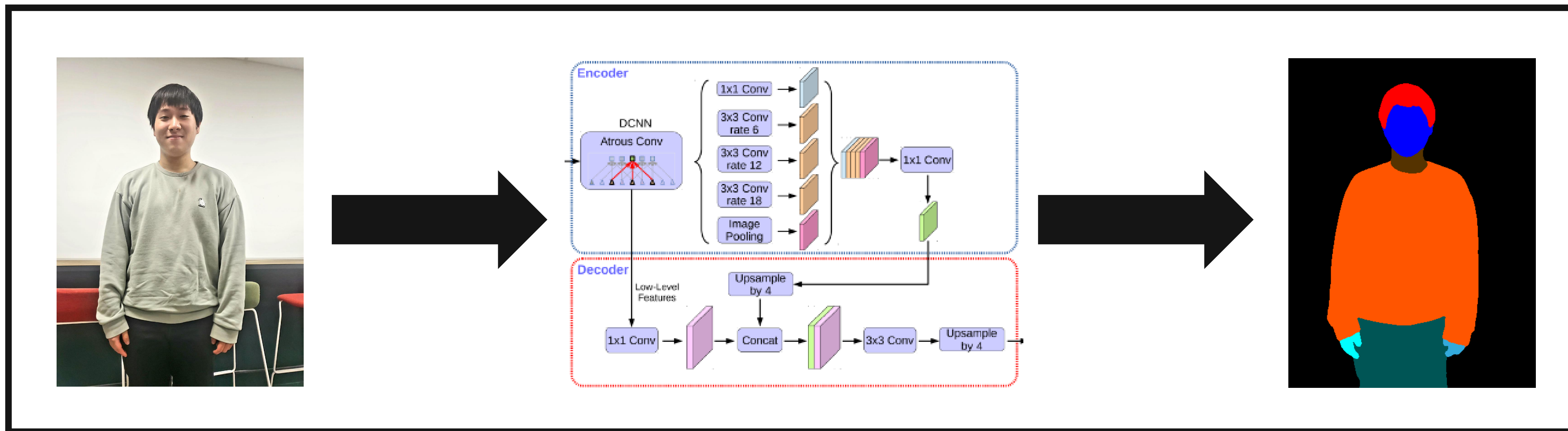
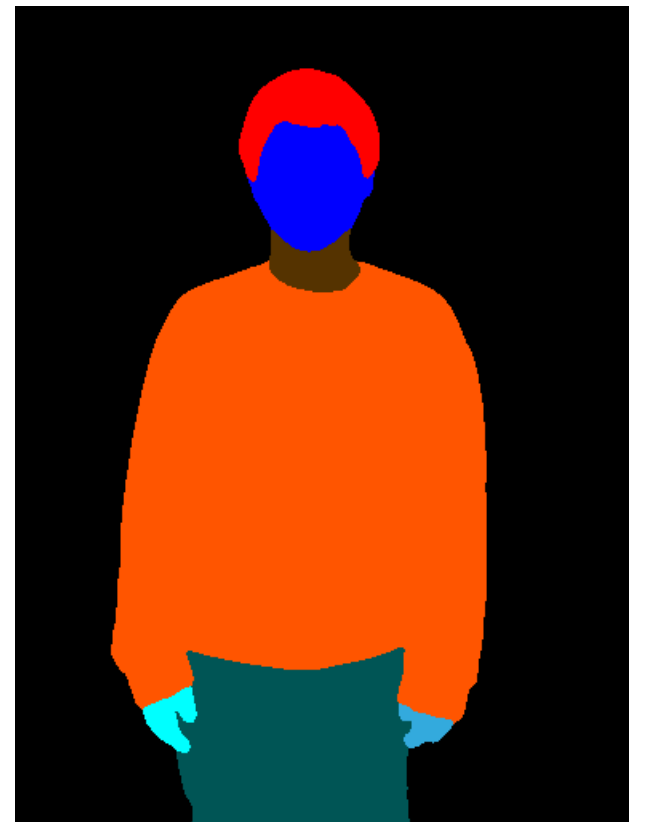
@sankalp2K

We use "densepose_rcnn_R_50_FPN_s1x" model in the [detectron2](#) repository :)

- 논문에서 preprocess시 사용했다고 답변한 Facebook research에서 제공하는 Detectron2 라이브러리를 사용해 Densepose Image를 추론함

Preprocess – Image 03, Segmentation Image

03



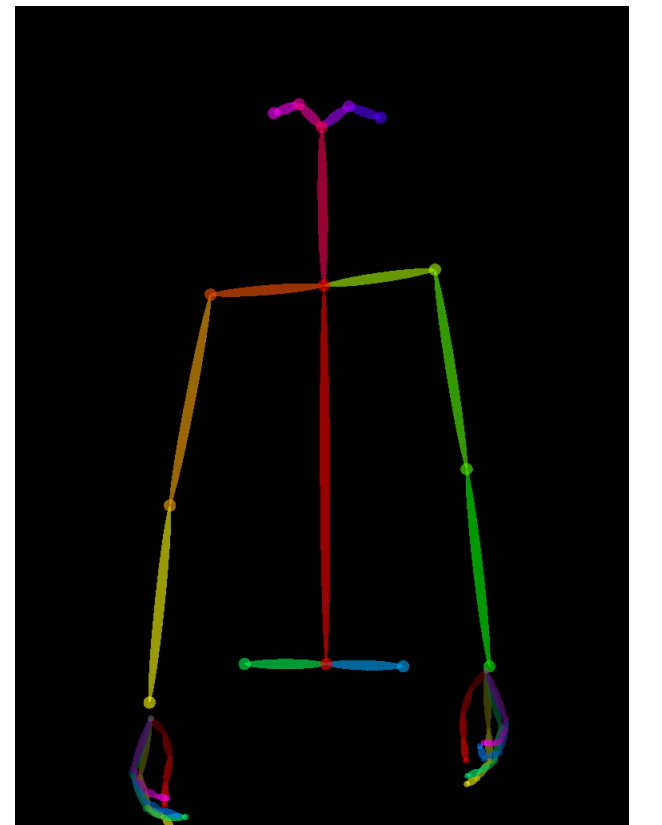
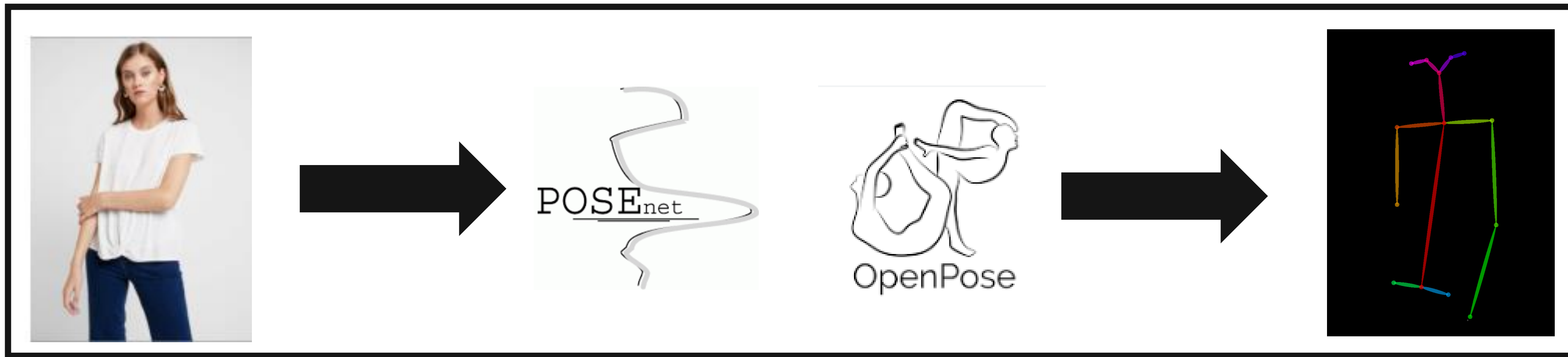
- DeeplabV3+ 기반 Semantic Segmentation 모델을 구축하고, 신체 부위 Semantic Segmentation 이미지 데이터셋을 사용해 학습하여 파이프라인에 추가함.
-> 현재는 Graphonomy 라이브러리의 pretrained 모델로 교체함.

```
Epoch 1/100
2500/2500 [=====] - 1204s 481ms/step - loss: 1.2589 - acc: 0.7594 - val_loss: 0.5215 - val_acc: 0.8243
Epoch 2/100
2500/2500 [=====] - 1095s 438ms/step - loss: 0.4329 - acc: 0.8613 - val_loss: 0.3193 - val_acc: 0.9177
Epoch 3/100
2500/2500 [=====] - 1095s 438ms/step - loss: 0.2714 - acc: 0.9240 - val_loss: 0.2143 - val_acc: 0.9365
Epoch 4/100
2500/2500 [=====] - 1094s 438ms/step - loss: 0.2226 - acc: 0.9360 - val_loss: 0.2050 - val_acc: 0.9417
Epoch 5/100
2500/2500 [=====] - 1093s 437ms/step - loss: 0.1959 - acc: 0.9431 - val_loss: 0.2027 - val_acc: 0.9424
Epoch 6/100
2500/2500 [=====] - 1093s 437ms/step - loss: 0.1930 - acc: 0.9440 - val_loss: 0.2273 - val_acc: 0.9449
```

Gong, K., Gao, Y., Liang, X., Shen, X., Wang, M., & Lin, L. (2019). Graphonomy: Universal human parsing via graph transfer learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 7450-7459).

Preprocess – Image 04, OpenPose coordinate

04



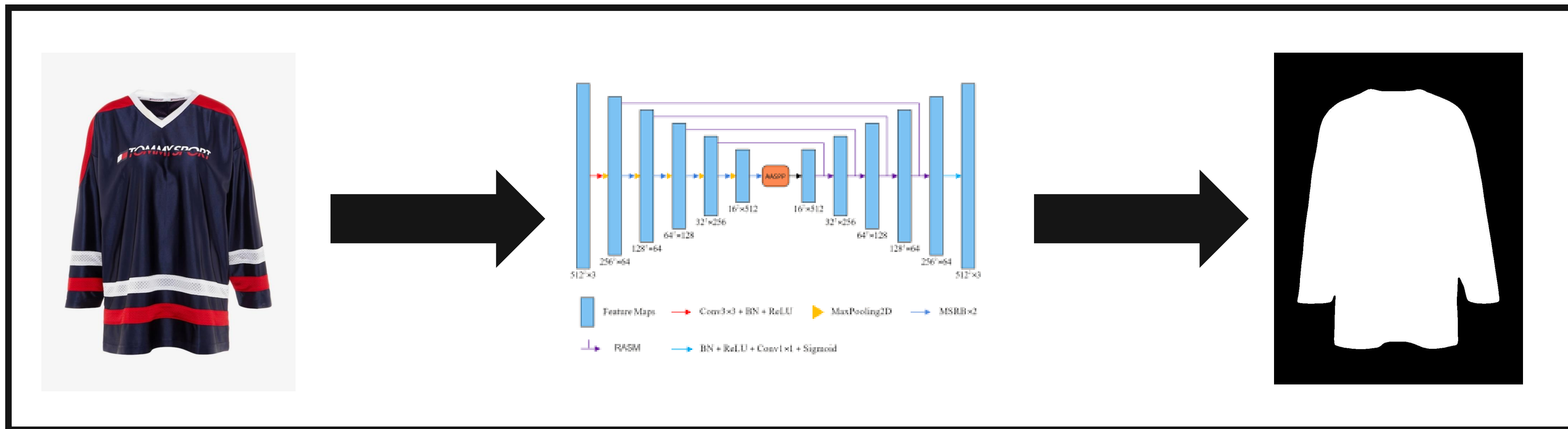
- Pose estimation을 가벼운 모델인 POSEnet으로 진행하고, 추후 OpenPose 좌표 형식으로 바꿔줌



- HR-VITON 내에서는 이렇게 생성된 openpose 좌표를 사용하여, 이미지의 상체 부분을 마스킹한 이미지와, 몸에 맞도록 비튼 옷 이미지를 생성함

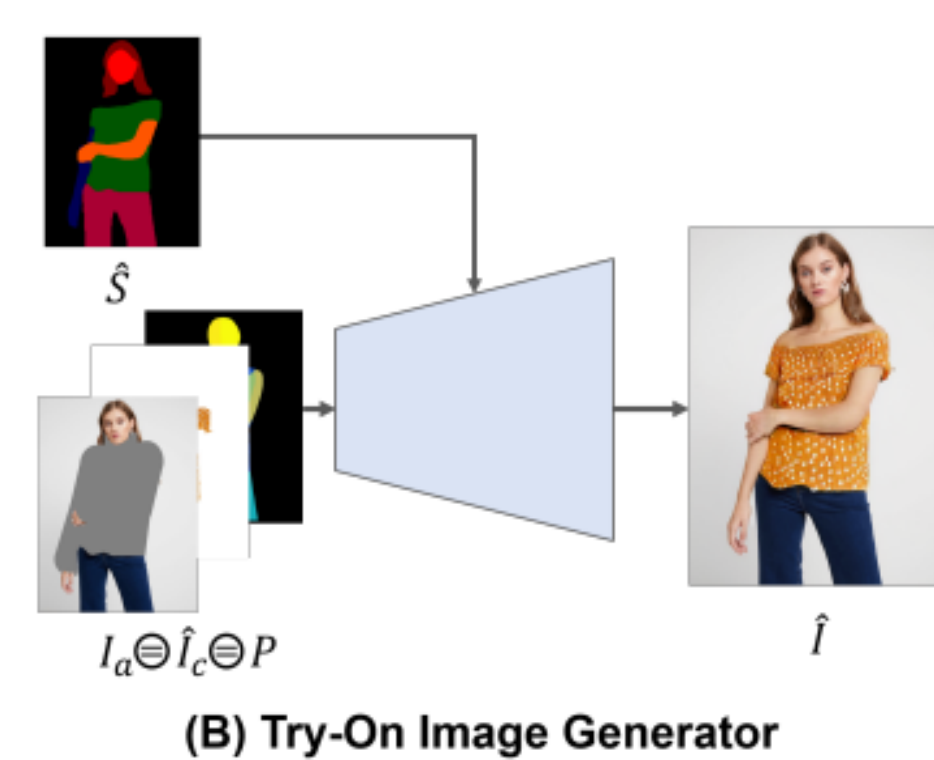
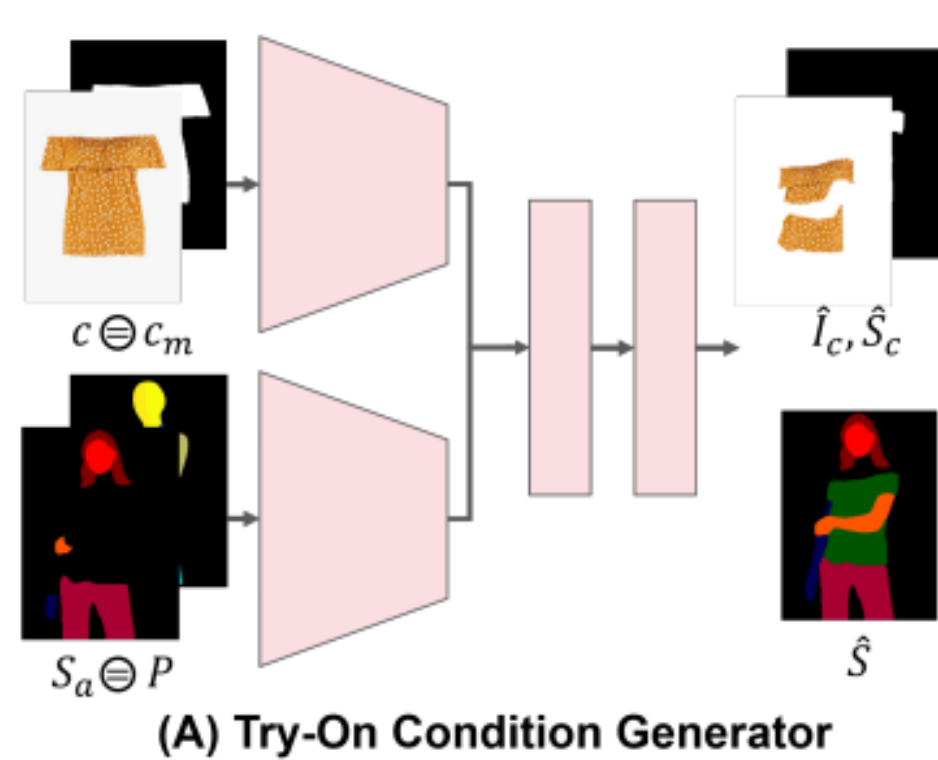
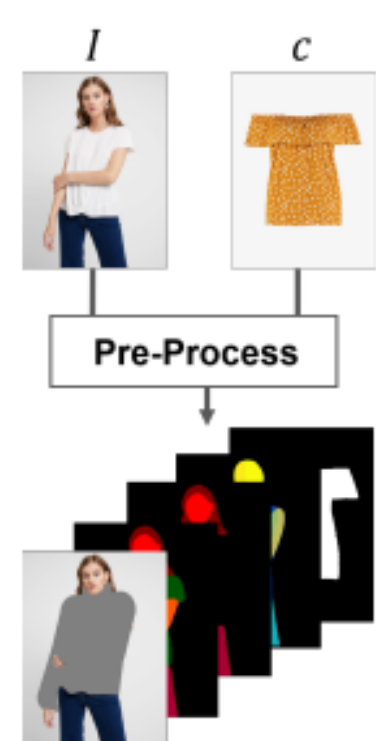
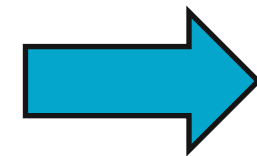
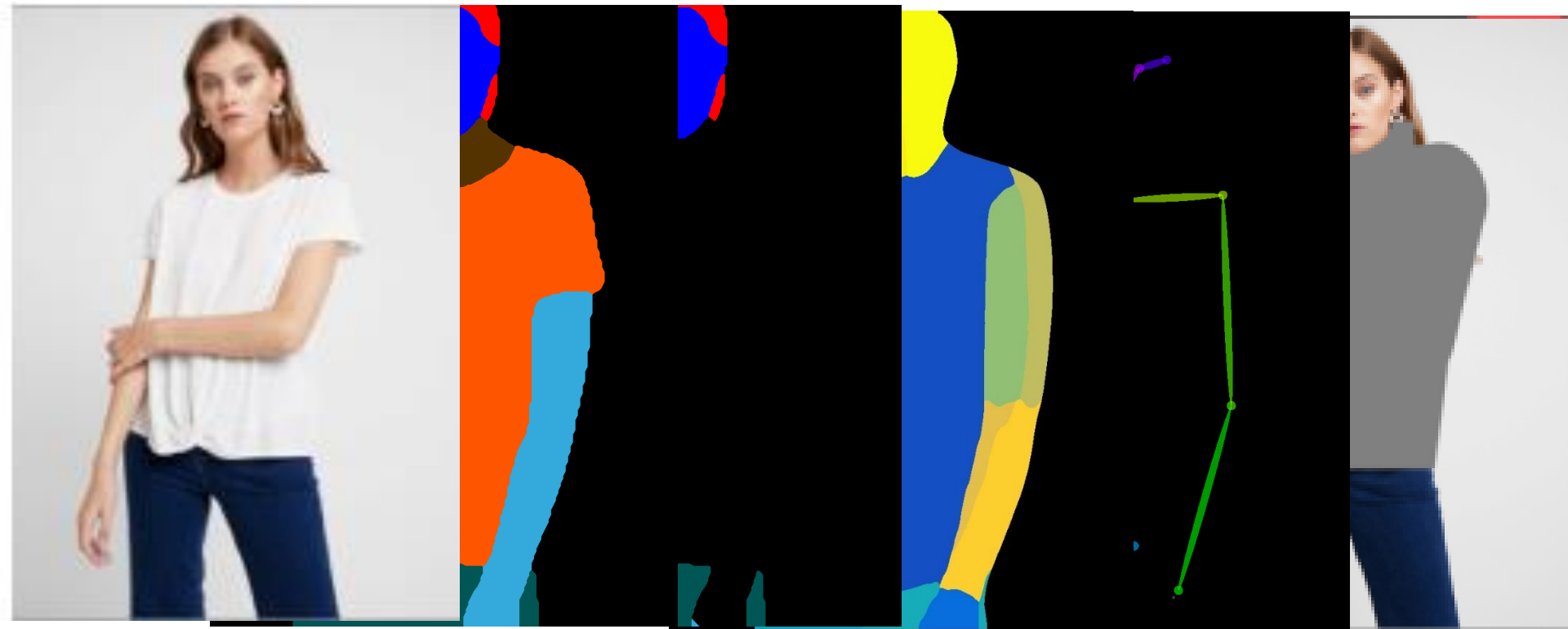
Preprocess – Image 05, Cloth Segmentation Image

05



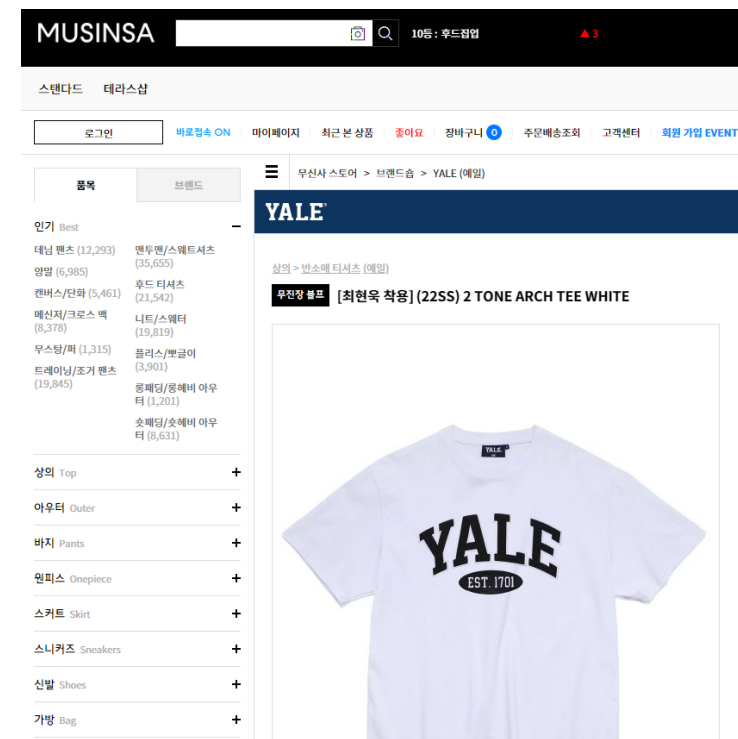
- Unet 기반 semantic segmentation 모델을 구축하고, 옷이 있는 부분과 없는 부분을 구분하는 Cloth Semantic Segmentation 모델을 학습하여 파이프라인에 추가함. -> 현재는 cloths_segmentation 라이브러리의 pretrained 모델로 교체함.

HR-VITON Inference



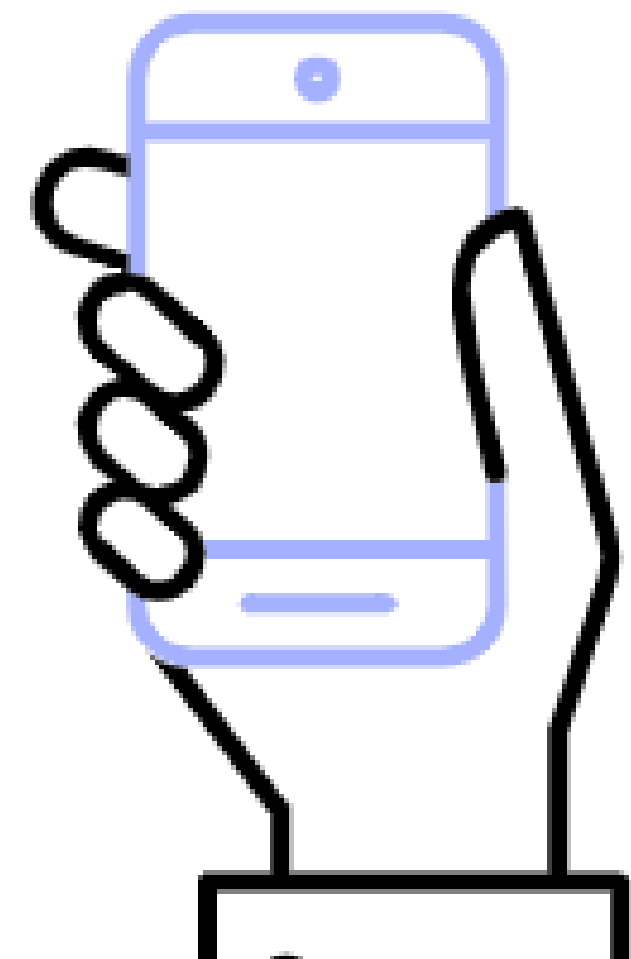
- 앞서 행한 preprocess 과정 모두 파이프라인으로 묶고, preprocess 과정을 통해 얻어낸 Input image들을 Base Model인 Pretrained HR-VITON 모델의 Input 이미지로 사용하여 Inference Image를 생성한다.

Architecture Overview

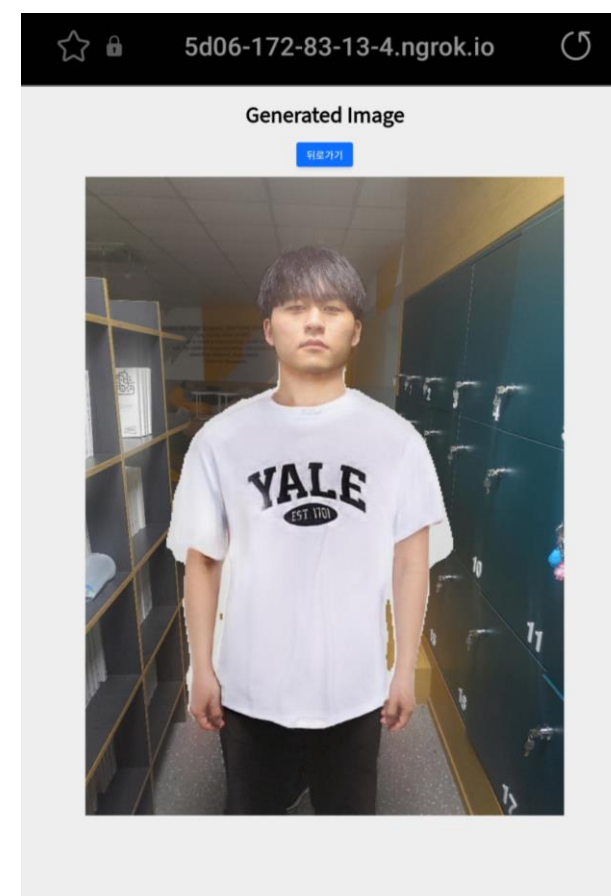


1. 자신의 사진 및
무신사에서 고른 옷
사진을 업로드한다.

2. Generate Image를 눌러
이미지를 생성한다.



User Device

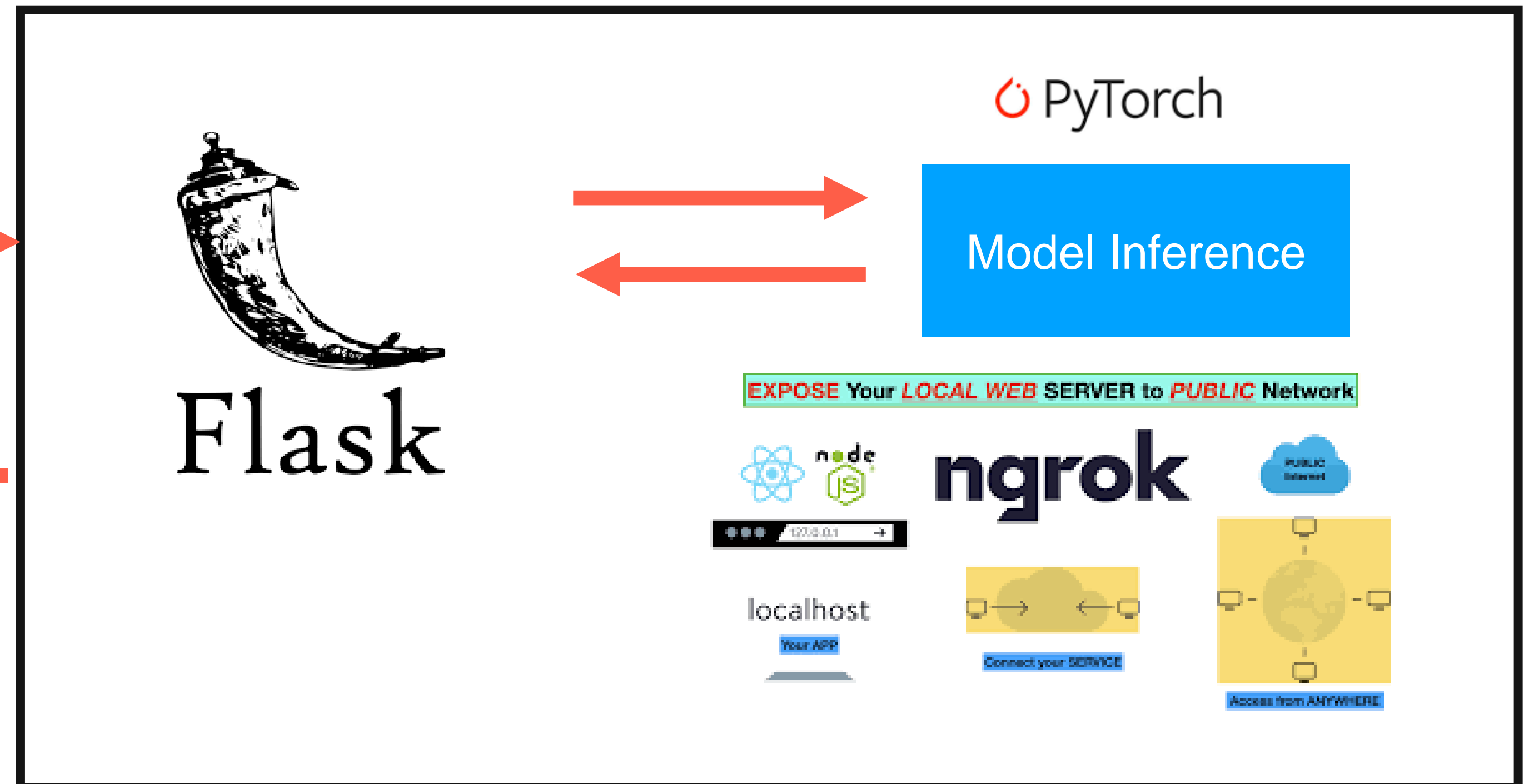


3. 약 15초간 기다린뒤,
생성된 이미지를 확인한다.

[Paperspace GPU Virtual Machine]
4 cpu, P5000 Vram 16GB



Paperspace



Inference Results



- 20201778 고재원



- 20211725 김민규

- Inference Result를 보면, 저희가 제안한 preprocess 파이프라인 및 모델 추가만으로 HR-VITON 논문 결과 재현에 성공했다고 할 수 있는 정도의 만족스러운 성능을 보이는 모습을 확인할 수 있습니다.

기대효과 및 서비스 시연

머신러닝
사이즈 추천

추천 사이즈
M(26-27)
* 2,312건의 사이즈 피드백 데이터 분석 결과

키 160cm | 몸무게 52kg [변경](#)

M(26-27) ▾ 사이즈가 고객님의게 잘 맞을 가능성

매우 낮음 낮음 보통 높음 매우 높음

- 사이즈 표준 체형에 키 90%, 몸무게 90% 일치하므로
- M(26-27)가 잘 맞을 가능성이 매우 높습니다.

Q. 비슷한 체형이 구매한 사이즈가 궁금하시나요? [바로 보기](#)

체형별
구매 사이즈 통계

M
비슷한 체형의 60% 고객들은 이 사이즈를 선택했습니다.
* 고객님과 유사한 체형의 구매 고객들의 데이터를 분석한 결과

키 160cm | 몸무게 52kg [변경](#)

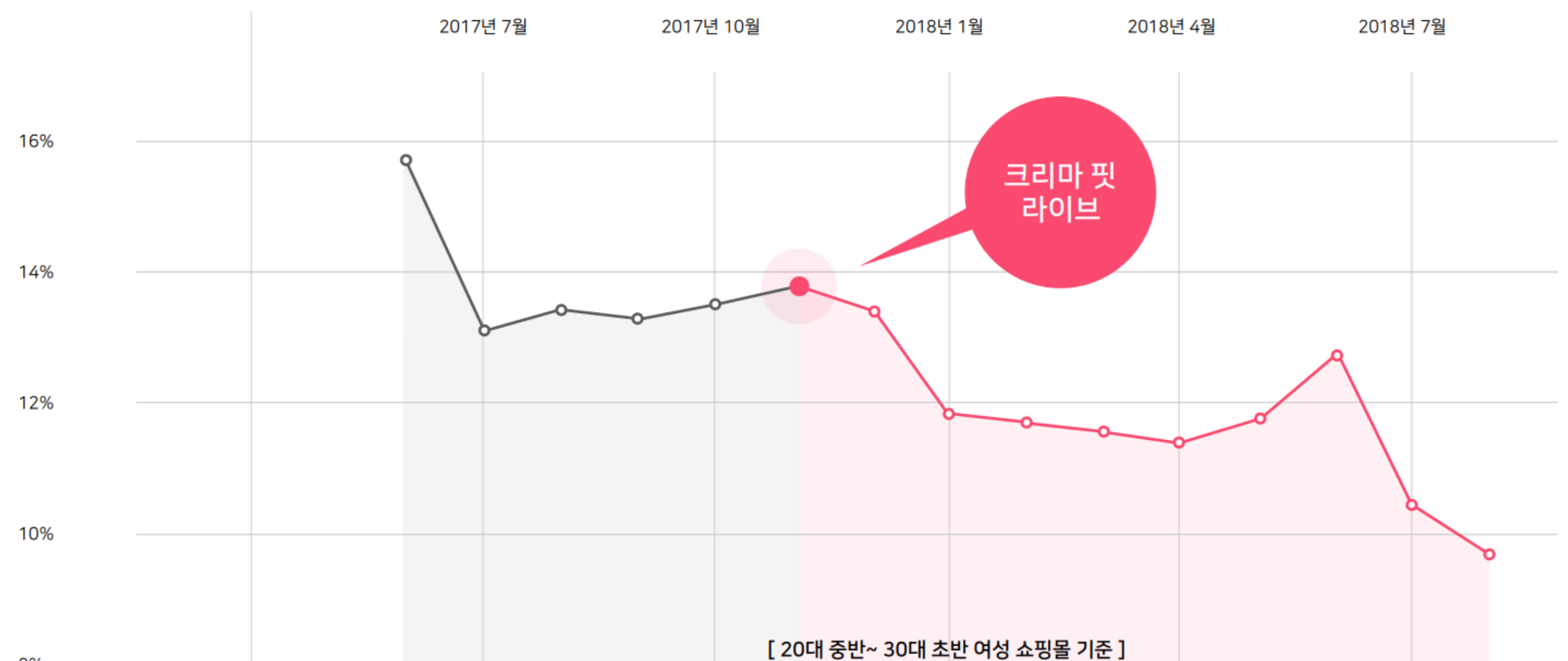
[1:1
M
S
Q. (

저희가 제안하는 모델을 사용하여
가상 피팅 서비스를 적용했을 경우의
기대효과로는

- 소비자의 만족 및 광고 효과
- 교환 / 반품률 감소 등이 있습니다

크리마 핏 사용 효과

실제 크리마 핏 서비스를 사용 중인 쇼핑몰의 경우
교환/반품률이 이전 대비 18% 감소하였습니다!



- 비슷하지만 조금 다른 서비스로는 온라인 옷 쇼핑몰에서 발생하는 문제를 머신러닝 AI로 해결하려고 등장한 "크리마" 솔루션이 있습니다
- 이 솔루션은 머신러닝 기술을 사용해 사용자의 몸무게, 키 등의 정보를 가지고 맞는 옷 사이즈를 추천해주는 간단한 머신러닝인데도, 효과적인 성능을 보이고 있습니다.



Thank you