

Visual Caster

모든 코드는 colab pro + 환경에서 작성 되었습니다.

```
| segmentation
|   | sample_image
|   |   | result : 샘플이미지 폴더에 있는 이미지들의 세그멘테이션 결과이미지
|   |   | *.jpg : 세그멘테이션 이전 이미지
|   | model_weight
|   |   | full_weight.h5 : 최종 서비스에 활용한 모델파일
|   |   | 타 학습에 사용했던 모델들은 파일의 크기를 고려하여 따로 첨부하지 않음
|   | data : segmentation에서 모델학습에 활용된 이미지들
|   |   | rawdata
|   |   | fashion : ICCV15 Fashion_Dataset(ATR)
|   |   |   | humanparsing : Instance-LevelHuman-Parsing
|   |   |   | full 데이터 링크 https://drive.google.com/uc?id=1B9A9UCJYMWTL4oBEo4RZfbMZMaZhKJaz
|   |   | 각 폴더에는 필요없는 데이터를 제거하고 큰 용량을 고려해 샘플이미지 10장씩 만을 저장해두었음
|   |
|   | ** 이 아래로 레이블링이 되어있는 ipynb 파일들은 모델을 학습 - 테스트하는데에 사용된 파일입니다.
|   | 주석은 겹치는 경우가 많기때문에 1. 에 해당하는 파일에만 주석을 달아 놓았습니다
|   | 0-1.test_DL_v3plus.ipynb
|   | 0-2.test_FCN.ipynb
|   | 0-3.test_Unet.ipynb
|   | 1-1. DLv3_base.ipynb
|   | 1-2.FCN.ipynb
|   | 1-3.Unet.ipynb
|   | 2-1. DL_v3+freezed_sgd0.01_train.ipynb
|   | 2-2. DL_v3+trainableTure_sgd0.001_train.ipynb
|   | 2-3. DL_v3+unfreeze_sgd0.01_train.ipynb
|   | 3-1. DL_v3+unfreeze_adam0.01_train.ipynb
|   | 4-1. DL_v3+unfreeze_sgd0.01+seblock_train.ipynb
|   | 5. Last_fulltrain.ipynb
|   | 6. model_visualization.ipynb
|   |   | model_visualization.ipynb : 모델의 결과 시각화
|
|   | django : python 내의 웹실행에 필요한 파일들을 담고있습니다.
|
|   | result : main 실행시에 스크래핑 파일들이 저장될 경로입니다.
|
|   | haarcascade_frontalface_alt.xml : facecrop에 사용된 파일 |
|   | arrange.py : 이미지들의 배치에 사용되는 함수들
|   | scrap.py : 데이터스크래핑에 사용되는 함수들
|   | seg.py : 학습된 모델을 사용하여 이미지를 forward시키고 후처리하는 함수들
|   | main.ipynb
```

- 서비스를 실행하는데에 사용된 파일입니다. 구글 코랩프로+를 기준으로 작성되어있으며 내부에 필요한 패키지 설치에 대한 코드 또한 포함합니다.
- 경로들은 오류를 방지하여 절대경로와 상대경로가 혼용되어있습니다. 절대경로의 경우 사용자의 설정에 맞게 변경하는 과정이 필요합니다.
- 데이터 스크래핑이 포함된 코드이기때문에 `headers` 딕셔너리 내의 `user-agent` 정보는 컴퓨터 환경에 맞게 변경하는 과정이 필요합니다.
- 셀의 마지막 링크를 누르면 구현한 웹화면으로 이동합니다

1. 아이디어 배경

1) 왜 '시각화'인가?

① 텍스트에서 이미지로의 정보전달 주체 변화

정보 전달에 있어 이미지의 사용이 급격하게 증가하고 있습니다. 이는 이미지가 지니고 있는 인지적 기능에 대한 관심의 증가와 관련이 있습니다. 이미지는 인간에게 이해와 소통의 수단으로 언어에 비해 정보 제공을 위한 노력이 더 많이 들어 열등한 것으로 간주되었으나 정보통신기술의 발달은 그동안 등한시되었던 이미지의 이해와 소통 기능을 다시 부활시키는 계기가 되었습니다.

이로 인하여 문자가 지배하는 텍스트 매체의 시대에서 이미지를 활용하는 매체가 주로 지배하는 시대로의 전환이 이루어졌고, 이에 사람들은 단순한 지식의 습득보다는 쉽게 느낄 수 있는 경험을 중시하게 되었습니다.

이는 일반적인 경험보다는 쉽게 경험할 수 없는 새로운 자극을 경험하기를 원하게 만드는 변화의 계기를 만들어 주었습니다. 의미와 감동을 체험하기를 원하며, 이성과 사고 중심에서 시각과 감성 중심으로 문화의 트렌드가 변화하고 있다는 것 입니다. 아래의 표를 참고하면, 세대가 넘어갈수록 텍스트보다는 이미지가 주인 SNS 플랫폼의 사용이 늘어나고 있다는 사실을 알 수 있습니다.

[표 2] 가장 많이 이용하는 SNS(1순위 기준)

(N: SNS 이용자(Z세대=84, 알파 세대=78), 단위: %)

SNS 이용 플랫폼 종류	Z세대	알파 세대
트위터	10.2	2.4
페이스북	31.0	33.7
싸이월드 미니홈피	3.9	0.0
카카오스토리	42.5	7.8
인스타그램	1.8	41.8
네이버 랜드	6.9	11.7
카카오 그룹	1.7	0.0
틱톡	-	2.5
기타	2.0	0.0

주: '틱톡'은 2021년 조사에서 신규 항목으로 조사됨에 따라 Z세대(2017년)에서는 조사내용 없음('-')

그림 1. Z 세대와 α세대 SNS 이용 플랫폼 비율

② 그림 우월성 효과와 이미지 인식 속도

정보가 시각적일수록 우리는 더 쉽게 인지하고 더 오래 기억하며, 이를 그림 우월성 효과라고 합니다. 1967년 발표한 저널의 내용에 의하면, 인지과학자인 Roger N Shepard(로저 뉴랜드 셰프드)는 피실험자에게 총 612장의 그림을 보여준 뒤, 재인지 검사를 진행하였습니다. 2시간 뒤의 검사에서 피실험자는 99.7%를, 2주 뒤에는 87%에 해당하는 그림을 기억하는 결과를 보였습니다. 또한 F.Pulvermüller(프리드만 풀버뮐러)와 MIT 연구에 따르면 텍스트의 인지시간은 90밀리 초 ~160밀리 초, 이미지의 인지시간은 13밀리초가 걸린다고 합니다. 연구 결과에 따르면 단순히 텍스트로 제공했을 때 보다 이미지로 제공하면 사용자 입장에서 훨씬 쉽고 빠르게 이해할 수 있습니다.

TABLE 1
PERCENT CORRECT RECOGNITIONS OF THE OLD STIMULUS IN TEST PAIRS FOLLOWING
INSPECTION SEQUENCES OF WORDS, SENTENCES, OR PICTURES

Experiment	No. Ss	Length of inspection series	Percent correct on test pairs			
			Mean	SD	Median	Inter-quartile range
I. Words	17	540	88.4	6.7	90.0	9.1
(old) (new)						
freq. freq.			82.1	10.7		
freq rare			86.7	11.5		
rare freq.			93.0	7.6		
rare rare			92.0	5.9		
II. Sentences (double deck)	17 2	612 1224	89.0 88.2	6.3 4.2	88.2 88.2	9.6 4.0
III. Pictures (after delay of)	34	612	96.7	3.8	98.5	4.4
2 hr	4		99.7	0.5	100.0	0.5
3 days	4		92.0	2.7	93.0	3.0
7 days	4		87.0	12.8	92.0	15.0
120 days	4		57.7	10.8	57.0	14.0

그림 2. Roger N Shepard(로저 뉴랜드 셰프드) 저널 실험 결과

2) 왜 '실시간 트렌트'인가?

① 이미지를 통한 실시간 트렌드의 이점

네이버에서 제공했던 실시간 검색어 서비스는 풍부한 정보 속에서 능동적으로 본인에게 필요한 정보를 소비하고 싶은 커다란 트렌드 변화에 맞춰 종료 되었다고 합니다. 하지만 이런 방향성은 확증편향과 선택적 인지를 일으킬 수 있는 일종의 정보 검열이 될 수 있다는 위험성이 존재합니다. 우리는 이미지를 이용한 새로운 방식으로 실시간 트렌드를 시각화함으로써 다시 실시간 트렌드의 순기능인 다양한 사용자들의 관심사라는 정보로서의 가치를 살리고자 합니다.

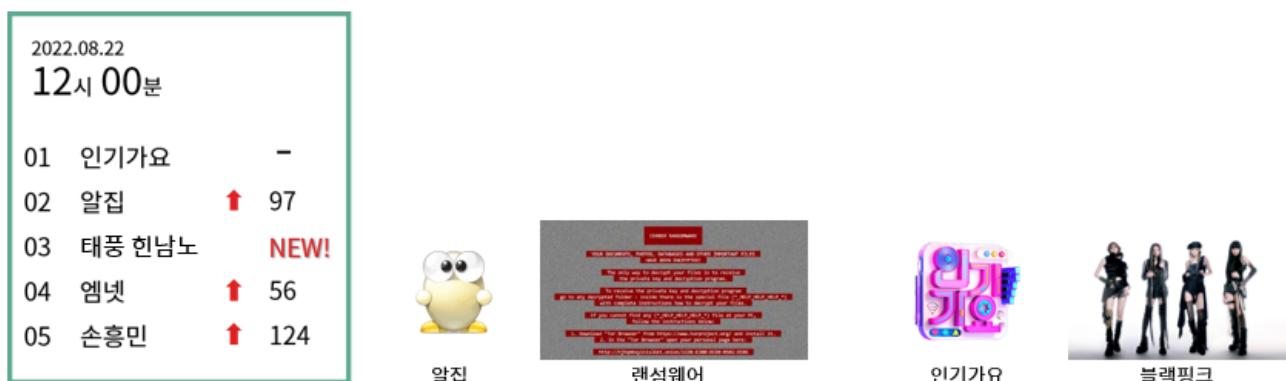
또한 단순히 텍스트가 아닌 이미지를 통하여 실시간 트렌드를 전달함으로써 정보의 객관성을 증가시키고, 이와 관련된 사용자 스스로의 생각을 불러일으킬 수 있을 것 입니다. 특정인물과 관련된 포털 내의 여론 경쟁을 보여주는 것이 아닌 그 인물의 이미지를 보여주는 것을 예로 들 수 있습니다. 'OOO 힘내세요', 'OOO 탈퇴하세요' 와 같이 주관성이 담긴 트렌드가 아닌 해당 인물에 대한 이미지는 트렌드를 보다 객관적으로 보여줍니다. 즉, 포털에서 제공하는 문자 정보를 그대로 받아들이는 것이 아닌 이 대상자와 관련된 이미지를 통하여 실시간 트렌드의 주체를 향한 자신의 주관에 편향을 일으키지 않고 다양한 정보의 수집을 촉진할 수 있습니다.



그림 3. 실시간 트렌드 이미지

② 관심사 이외의 정보 확장

실시간 트렌드를 보는 사용자들은 모두 다른 지식과 정보를 가지고 있으며 관심 분야 또한 모두 다릅니다. 트렌드를 텍스트만으로 제공한다면 관련 지식이나 관심이 있는 사용자는 단어만을 통해 인지할 수 있지만, 그 외의 사람들은 그 단어가 무엇을 의미하는지 알지 못합니다. 하지만 이를 이미지를 통해 제공한다면 관심사나 지식 수준이 다른 사람들도 쉽게 해당 트렌드에 대해서 파악이 가능하며 폭넓은 이해로 이어질 수 있습니다. 키워드란 트렌드 그 자체이거나 대상이 명확하지 않은 경우 여러가지 검색 기사 타이틀의 유사도 검색을 통해 뽑아낸 객체 이미지를 찾을 때 사용할 단어이자 트렌드가 실시간 이슈 트렌드에 오른 이유를 설명하는 단어입니다. 예를 들어, 알집이 실시간 트렌드에 올라왔을 때, 단어만으로는 이것이 이슈가 된 이유를 파악할 수 없지만, 이미지로 랜섬웨어가 나오게 되면 알집의 이슈 이유가 랜섬웨어와 관련되었다는 것을 파악할 수 있습니다. 또한 처음 들어보는 이름의 드라마가 실시간 검색어에 올라왔을 때, 단어만으로는 이것이 드라마인지 파악할 수 없지만, 이미지로 드라마 포스터가 올라오고, 이미지의 확장으로 주연 배우나 조연 배우의 이미지까지 나타내진다면 이 단어가 드라마를 나타내며 주연이 누구인가까지 해당 단어에 대해 폭 넓은 파악이 가능합니다.



2. 핵심기능

서비스의 핵심 기능은 실시간 트렌드의 이미지 시각화 및 Layout, 그리고 함께 제공되는 기사 요약 카드뉴스 두 가지 부분으로 나눌 수 있습니다.

1) 요약 카드 뉴스 생성

실시간 트렌드에 대한 기사들을 크롤링을 통해 가져와 제목의 문장 간 유사도를 측정하여 높은 기사를 선택합니다. Abstractive 방식을 이용하여 크롤링한 기사의 원문을 요약하고 선택한 기사의 제목과 이미지, 기사를 요약한 요약본을 통해 카드뉴스를 생성합니다.

2) 이미지 Layout 생성

실시간 트렌드 크롤링 시 트렌드가 개체명을 나타내는 명사인 경우 해당 트렌드에 대한 이미지를, 아닌 경우에는 기사 제목에서의 개체명을 가져와 각 키워드에 해당하는 이미지를 도출합니다. 트렌드가 인물인 경우 segmentation 모델을 통해 배경을 제거한 인물의 이미지, 기간이나 드라마 등은 해당하는 로고 이미지를 가져옵니다. 가져온 이미지를 배치 모델을 통해 Layout 이미지를 생성합니다.

생성된 2개의 결과물을 통하여 웹 서비스를 제공하여 이미지를 띄운 뒤, 이미지에 마우스 오버를 하게 되면 팝업으로 랭킹, 트렌드, 키워드를 보여주고 그 이미지에 해당하는 요약된 카드 뉴스를 볼 수 있고 클릭 시 해당 뉴스의 페이지로 이동합니다.



그림 3. 서비스 기능(이미지 Layout + 요약 카드뉴스)

3. 사용데이터

실시간 트렌드에 주로 올라오는 트렌드의 종류는 인물관련 트렌드이기에 이를 이미지로 표현하기에 앞서 segmentation을 진행합니다.

1) Instance-Level-Human-Parsing



- 라벨링 된 2인 이상 이미지 33279장 ### 2) ICCV15_fashion_dataset(ATR)



- 라벨링 된 1인의 전신 이미지 17706장

4. 학습데이터의 품질검증 및 품질보완

1) 학습데이터의 품질검증



Raw 데이터의 목적은 사람이 착용하는 의류별로 segmentation하는 것이었으므로 액세서리를 제외하고 구별, 마스킹되어 있습니다. 위 특징은 Raw 데이터를 그대로 모델에 학습할 시에는 모델이 예측함에 있어 목걸이 시계 같은 액세서리를 비운 채로 예측할 수 있다는 문제점을 가지게 됩니다.

2) 학습데이터의 품질보완



① – Raw 데이터의 마스킹을 전처리한 후 학습

첫번째 방안으로는 Fill hole 알고리즘을 활용하여 마스킹되지 않은 악세서리와 인간이 착용한 사물들을 직접 마스킹한 뒤, 다시 학습을 시키는 것으로 마스크를 전처리해 주었습니다.

데이터셋의 모든 마스크에 대하여 다음과 같은 전처리를 해주고 모델의 성능을 다시 테스트해본 결과, 원하던 방향인 귀걸이와 목걸이를 함께 segmentation하는 것에는 성공하였으나, 그 외의 인물의 외곽선이 명확하지 않고 배경에 보이는 사물에 대하여 인물로 인식해 버리는 문제가 발생하였습니다.

이는 전처리를 하는 과정에서 어디까지를 사람으로 할 것인가에 대한 기준이 모호함에서 오는 문제로, 단순한 장신구는 문제가 없었지만, 사람이 착용한 물건이 신체의 안쪽에 있는지, 혹은 바깥으로 향하게 배경 쪽에 위치한지에 대하여 구분하고 이를 마스크하는 데 있어 정확한 판단 기준을 잡기 어려웠기 때문입니다.

또한 사람이 직접 라벨링을 한 번 더 거치는 과정이 있어 시간적 한계를 가질 수 밖에 없다는 문제를 가지고 있었습니다.

② – Raw 데이터로 학습을 한 후 예측결과를 후처리



따라서 두번째로는 악세서리가 비어있는 원데이터로 학습을 한 뒤 예측을 하는 마스크의 비어있는 부분을 Fill hole 알고리즘을 활용하여 채워 넣는 방식으로 진행하였습니다. 이는 인력이 많이 소모되지 않을 뿐만 아니라 이상적인 결과를 보여주었습니다.

이에 Raw 데이터의 문제점에 대한 해결방안으로 2 – Raw 데이터로 학습을 한 후 예측결과를 후처리를 선택하였습니다.

5. Data Augmentation

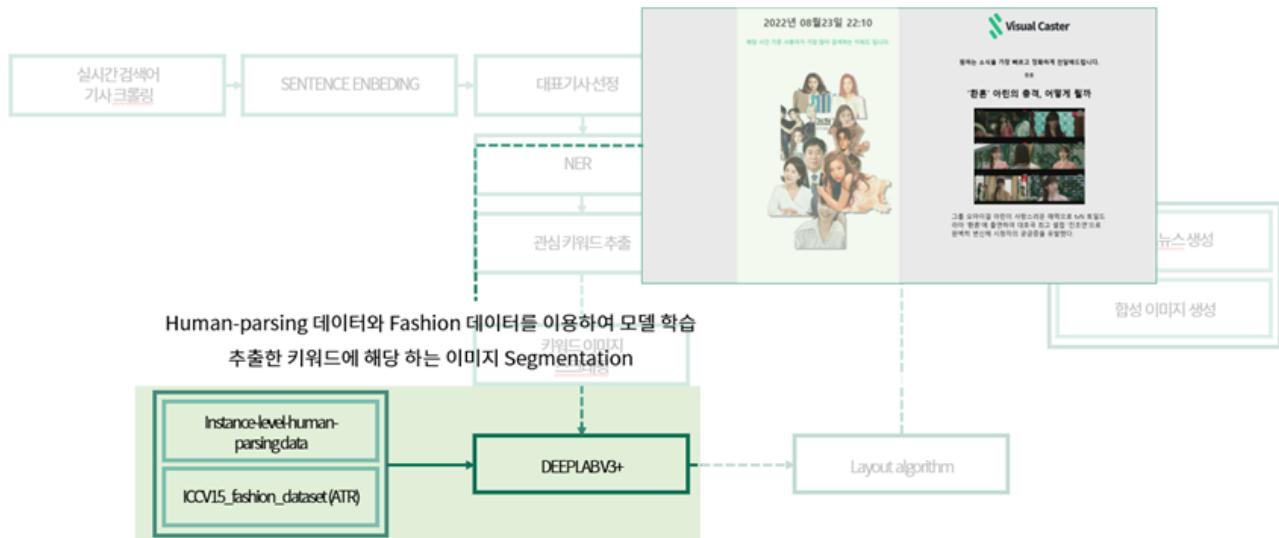


데이터의 다양화를 위해 매 epoch마다 원데이터를 조금씩 변형하여 Data Augmentation을 실시하였습니다.

모든 요인을 확률적으로 적용하였으며 이미지와 마스크 모두 같은 변환을 수행하였습니다.

1) crop - 512x512 이미지를 400x400과 440x440 사이의 크기로 crop해준 뒤 모델의 입력 사이즈인 512x512 resize해주었습니다. 2) 밝기 조절 - 밝기 조절을 통해 임의의 밝기로 바꾸어 주었습니다. 3) 채도 조절 - 채도 요인을 0.2로 설정하여 채도를 조절하였습니다. 4) 반전 - 이미지를 50%의 확률로 좌우 반전을 하였으며, 상하반전은 하지 않았습니다. 5) 회전 - 이미지를 -15도에서 15도 사이에서 임의로 회전을 시켜주었습니다.

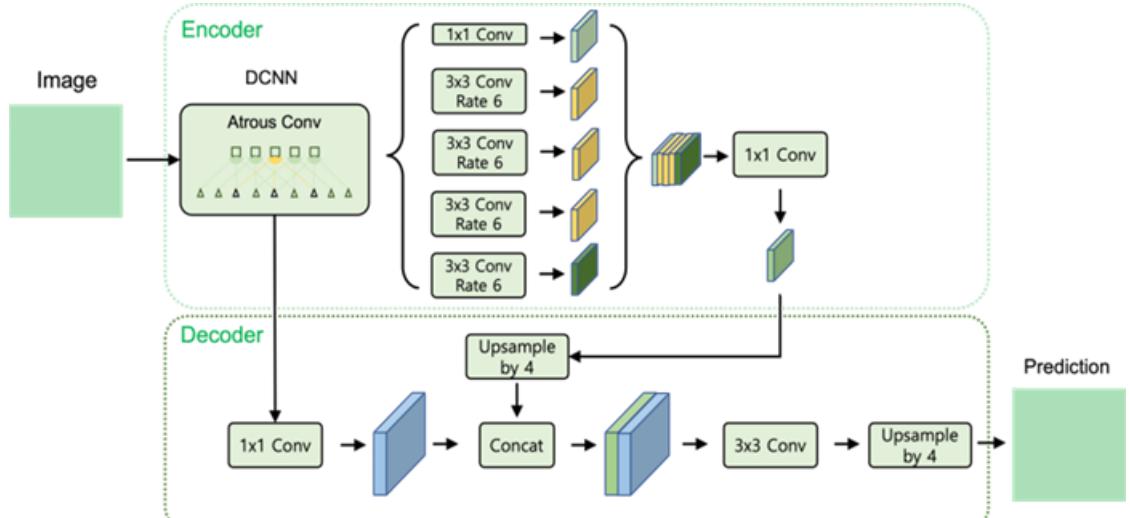
서비스 제공과정



6. Deeplab V3+

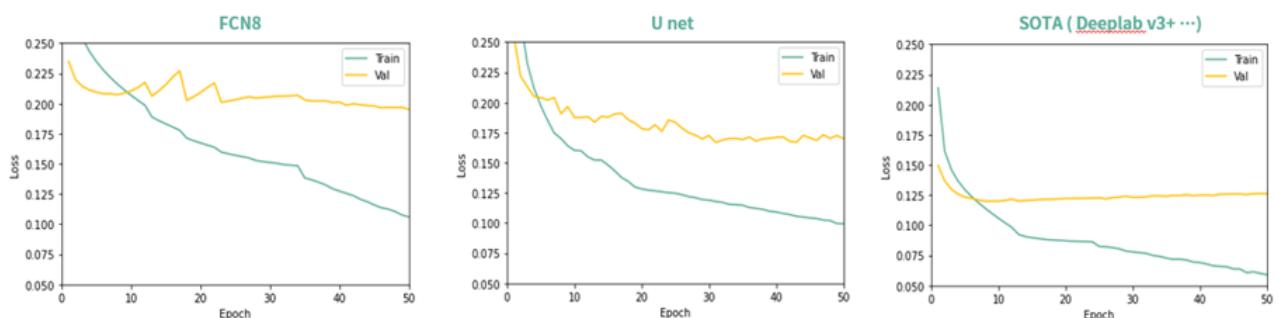
실시간 검색어의 주를 이루는 키워드는 인물이기에 인물을 segmentation 후 이미지를 배치하여 보여주고자 하였습니다. 이에 여러 실험을 통해 segmentation 모델인 Deeplab v3+를 선택하였고 간단한 실험을 통하여 모델을 트레이닝 시켰습니다.

DeeplabV3+에서는 여러 rate의 Atrous conv을 사용해 다양한 크기의 receptive field를 확인하는 ASPP(Atrous Spatial Pyramid Pooling)를 활용하였습니다. 그리고 인코더 디코더 구조를 차용함으로써 ASPP로 추출한 피처맵을 다시 확장하여 픽셀단위의 개체 예측



을 실시하였습니다.

1) Segmentation 모델의 선택



Segmentation 모델의 구축을 위해 FCN8과 U net, Deeplab v3+ 을 비교해보았습니다. Train 셋과 validation 셋에 대하여 loss를 비교해본 결과, FCN8과 U net에 비해 Deeplab v3+가 이상적으로 줄어드어 채택하게 되었습니다.

2) backbone 모델의 선택

Model	Acc	Params
Xception	79.0%	22.9M
ResNet152V2	78.0%	60.4M
EfficientNetV2S	83.9%	21.6M
EfficientNetV2M	85.3%	54.4M
EfficientNetV2L	85.7%	119.0M

일반적인 deeplab v3 plus 에서는 backbone으로 res-net / xception 모델을 주로 이용합니다.

본 조는 그 보다 더 발전한 CNN모델인 EfficientNet을 backbone으로 변경하였습니다. 실시간으로 세그멘테이션이 이루어지는 본 프로젝트에 특성상 forward 시간이 오래 걸리는 큰 모델은 사용하지 못했습니다. 이에 몇 가지 실험을 통해 어느 정도의 성능을 가지고 적절한 속도를 가지는 EfficientNetV2M의 image net pretrained model을 backbone으로 사용하여 transfer learning 했습니다.

각 레이어의 Feature map의 크기가 맞는 위치의 레이어를 추출하여 deeplabv3+의 레이어로 사용하도록 했습니다.

3) backbone 모델의 파인튜닝 여부

앞 단의 backbone 모델을 튜닝할 것인지, 튜닝하지 않고 그대로 사용할지에 대한 실험을 실시하였습니다.

실험시에는 sgd optimizer를 사용하였고 learning rate decay를 적용한 상태로 실험을 실시하였습니다.

① backbone freeze + learning rate 0.01

backbone 모델을 학습이 불가능한 상태로 고정을 하고 backbone에서 미리 학습된 필터와 feature map을 그대로 사용하도록 하였습니다. 대신 ASPP와 그 이후의 conv layer들이 빠르게 학습될 수 있도록 learning rate을 0.01로 설정하여 학습하였습니다.

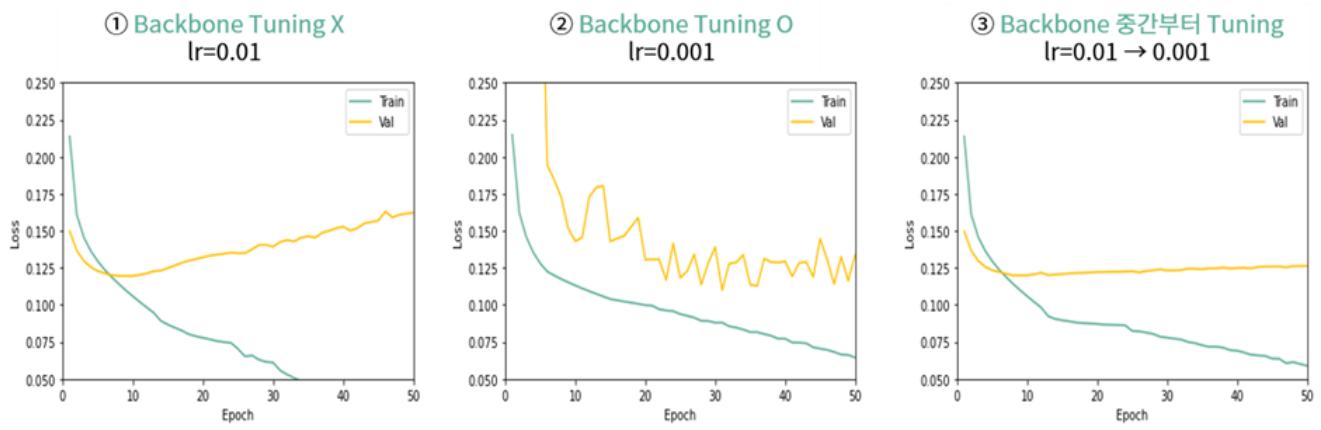
② backbone unfreeze + learning rate 0.001

처음부터 backbone을 학습이 가능한 상태로 두고 학습을 하였습니다. pre-train 된 layer들이 과도하게 조정되어 pre-train의 의미를 잃지 않도록 learning rate를 0.001로 두어 학습하였습니다.

③ backbone freeze + learning rate 0.01 -> 0.001

①, ② 두가지 개념을 모두 적용하였습니다. 처음에는 backbone을 학습 불가능하게 하고 타layers를 빠르게 학습되게끔 learning rate을 0.01로 적용하여 학습하였습니다.

학습도중 loss가 정해 놓은 epoch 동안 줄어들지 않으면, 학습이 불가능했던 backbone모델의 layers를 학습 가능하게 변경하고 learning rate을 0.001로 조정하여 학습하였습니다.



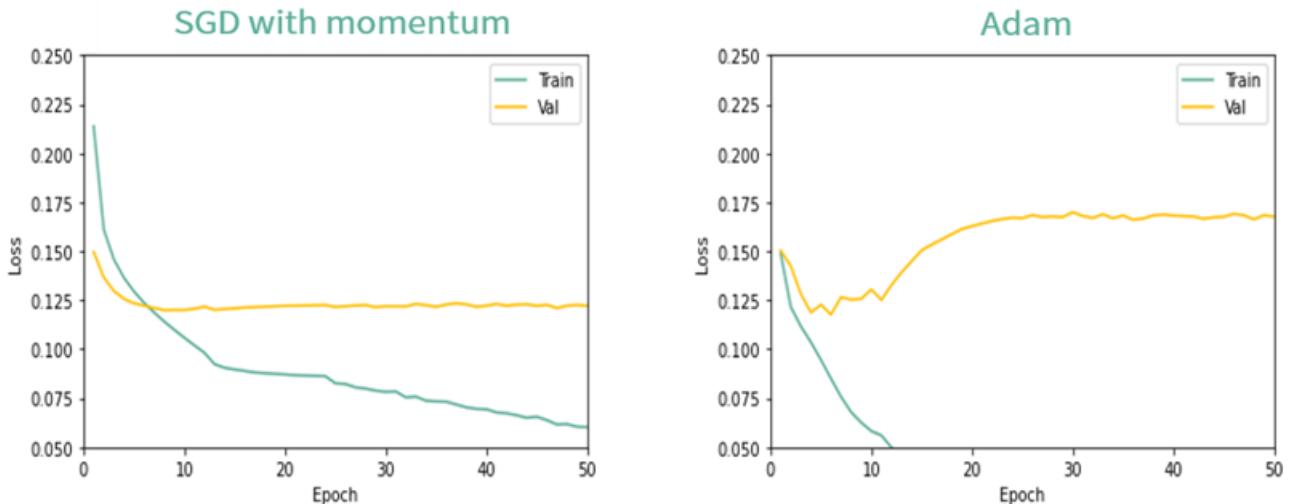
50000개의 이미지 중 5000개를 val set으로 활용하여 실험을 실시한 결과

①의 경우 빠르게 수렴하나 학습을 거듭할수록 val loss가 증가하는 과적합이 빠르게 일어나는 것을 확인하였고 ②의 경우 backbone val loss가 크게 흔들리는 것을 확인하였습니다. 그에 반해 ③의 경우는 val loss가 조금씩 증가하고 있지만 그 정도가 크지 않고 train loss 또한 안정적으로 줄고 있는 모습을 보였습니다. 이 실험 결과 ③의 방법을 택하여 모델을 학습하기로 결정하였습니다.

4) optimizer tuning

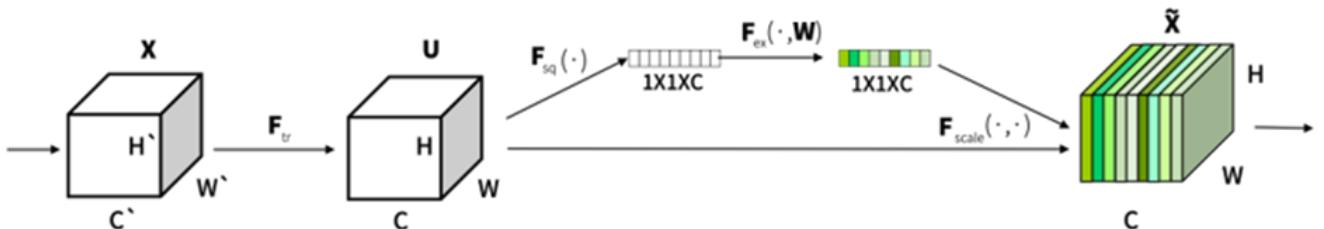
optimizer는 보편적으로 사용하는 Adam과 momentum을 적용한 sgd를 같이 비교하였습니다.

두 optimizer 모두 앞에서 설명한 val loss가 줄어들지 않으면 backbone을 unfreezing 하는 방법으로 learning rate을 0.01에서 0.001로 적용되게 하였고 learning rate decay를 적용한 상태로 실험을 실시하였습니다.



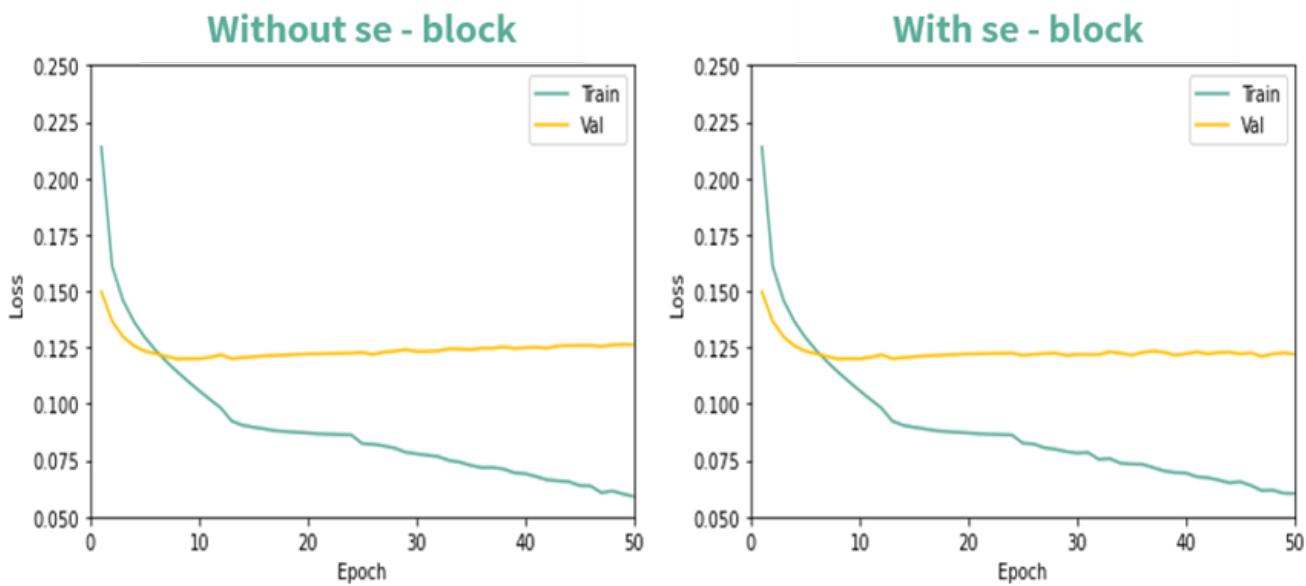
Adam의 경우 빠르게 학습되지만 과적합이 크게 발생하는 것을 확인할 수 있었고 sgd의 경우 val loss가 안정적으로 줄어들었습니다. 이에 momentum을 적용한 sgd를 optimizer로 선택하였습니다.

5) SE-block



SE-block은 채널 간의 상호작용을 학습한 뒤, 그 정보를 사용해 채널 단위로 가중치를 학습하여 성능 향상을 이끌어냈습니다.

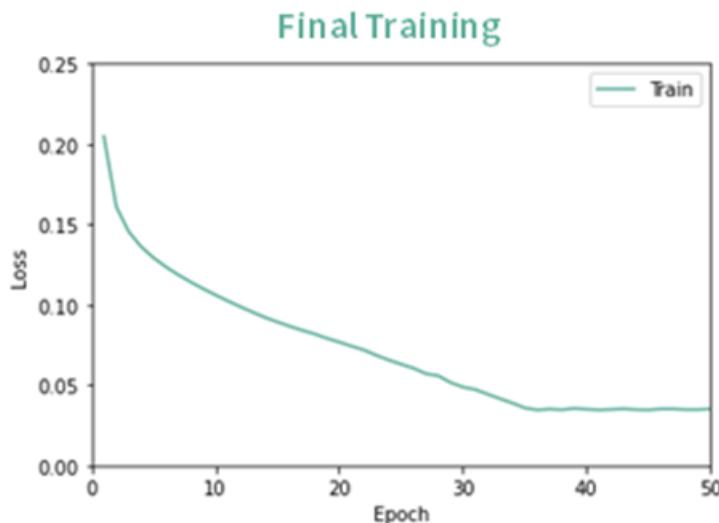
이는 자연어의 attention 개념이라고 볼 수도 있습니다. SE-block은 유연하고 다양하게 적용 가능하고 많은 추가 연산 양 없이 CNN 모델의 성능을 이끌어 낼 수 있는 방법입니다. SE-block의 개념을 본 모델의 ASPP 이후에 적용하여 성능의 향상을 꾀하였습니다.



근소한 차이지만 SE-block을 적용한 것이 성능이 좀 더 좋았습니다. (best val loss 0.1199 / 0.1193) 이에 최종모델에도 SE-block을 적용하였습니다.

6) 최종모델

위의 실험결과를 반영하여 50000개의 이미지를 모두 trainset 학습하였습니다.



이와 같은 실험 결과들을 반영하여 최종 모델 학습에 있어서 Deeplab V3 +와 Efficient netV2M를 사용하였고, sgd 옵티마이저를 적용하며 백본을 초반에 프리징 하였다 후에 언프리징 하는 방식을 채택하였습니다. 또한, SE-block을 추가 적용하여 학습하였습니다.

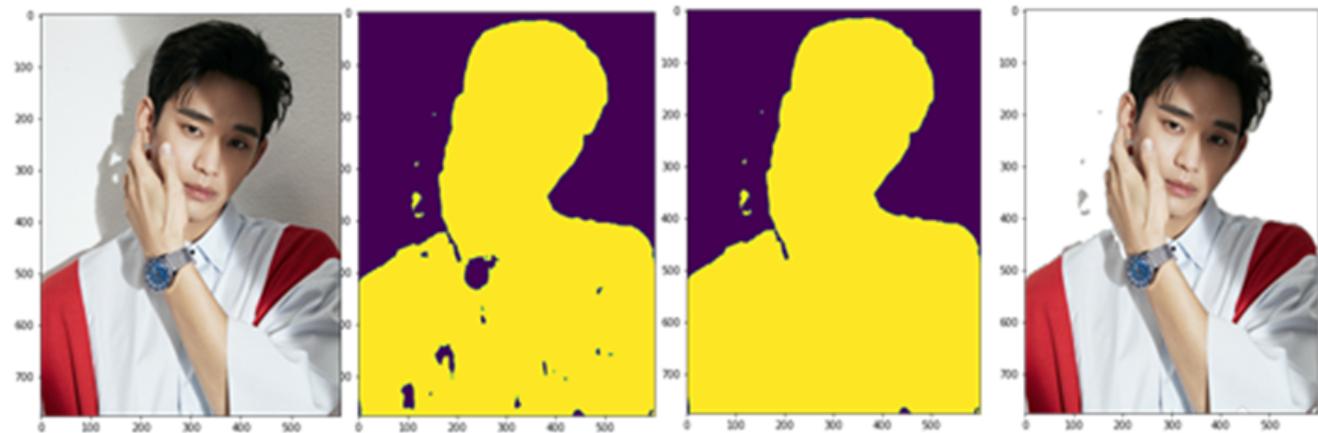
7) 결과 시각화



실험 결과를 반영한 최종 모델을 통한 Segmentation의 결과를 비교해 보았을 때 모델이 안정적으로 학습된 결과를 확인할 수 있었습니다. 과한 패턴이 옷에 있거나 배경과 옷의 색이 명확히 구분되지 않는 경우를 제외한다면 인물을 훌륭하게 추출하는 것을 확인할 수 있습니다.

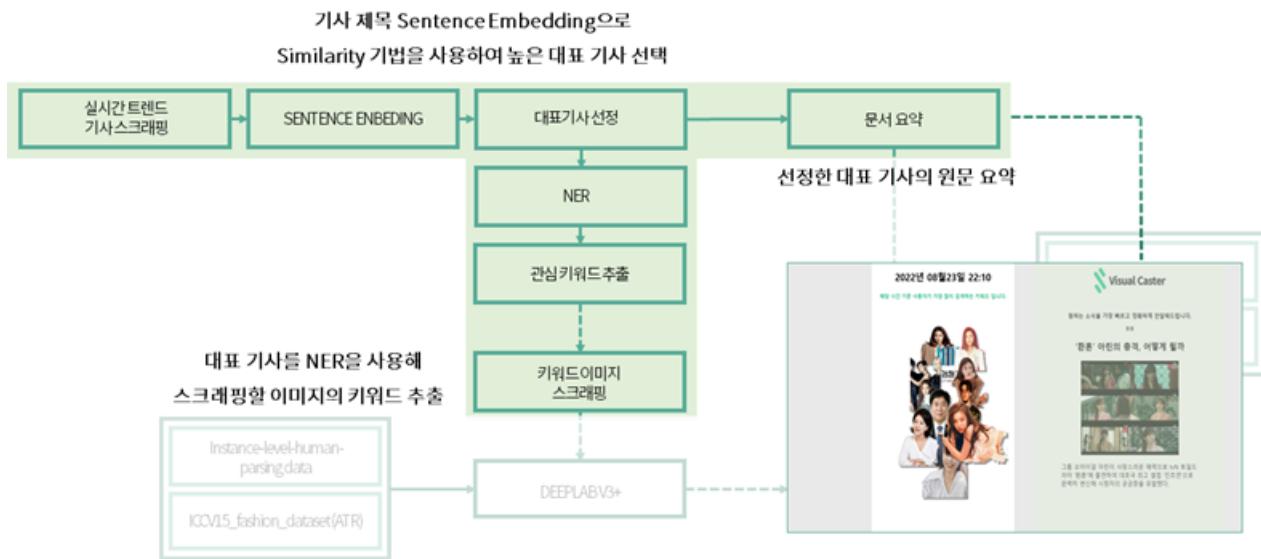
8) 결과 후처리

데이터 수집에서 언급하였듯이 수집한 데이터에 한계가 있었고 이에 모델을 통과한 결과에 약간의 후처리를 실시하였습니다. 데이터 셋에서 액세서리를 제외한 마스킹을 제공하기 때문에 후처리를 하지 않은 결과물에서는 시계부분을 마스킹하지 못하는 것을 확인할 수 있습니다. 이에 cv2의 find contour를 적용하여 contour안의 contour, 즉 마스킹 안에 구멍처럼 비워진 부분을 채우도록 하여 후처리하였습니다.



위의 이미지와 같이 오른쪽에서 시계를 마스킹하지 못하였던 것을 후처리를 통하여 보완해주었습니다.

7. 이슈 키워드 추출 및 기사요약



해당 과정에서는 카카오 브레인에서 다양한 자연어 태스크에 대응 가능한 통합된 형태의 오픈소스 자연어 프레임워크 PORORO(Platform Of neuRal mOdels for natuRal language processing)를 사용하였습니다.

1) 네이버 시그널 검색어 스크래핑

시그널 실시간 검색어
현재 기준 사용자가 가장 많이 검색 하는 키워드입니다.

2022년 9월 2일 금요일 오전 11:19

1 힌남노 경로	-	6 이찬원	-
2 8월 소비자물가 5.7% ↑	-	7 추가 확산 가능성	-
3 허위경력 거짓 해명	-	8 박셀바이오	-
4 량현 2년째 연락두절	-	9 입국 전 해외여행	-
5 제니	-	10 나지완	-

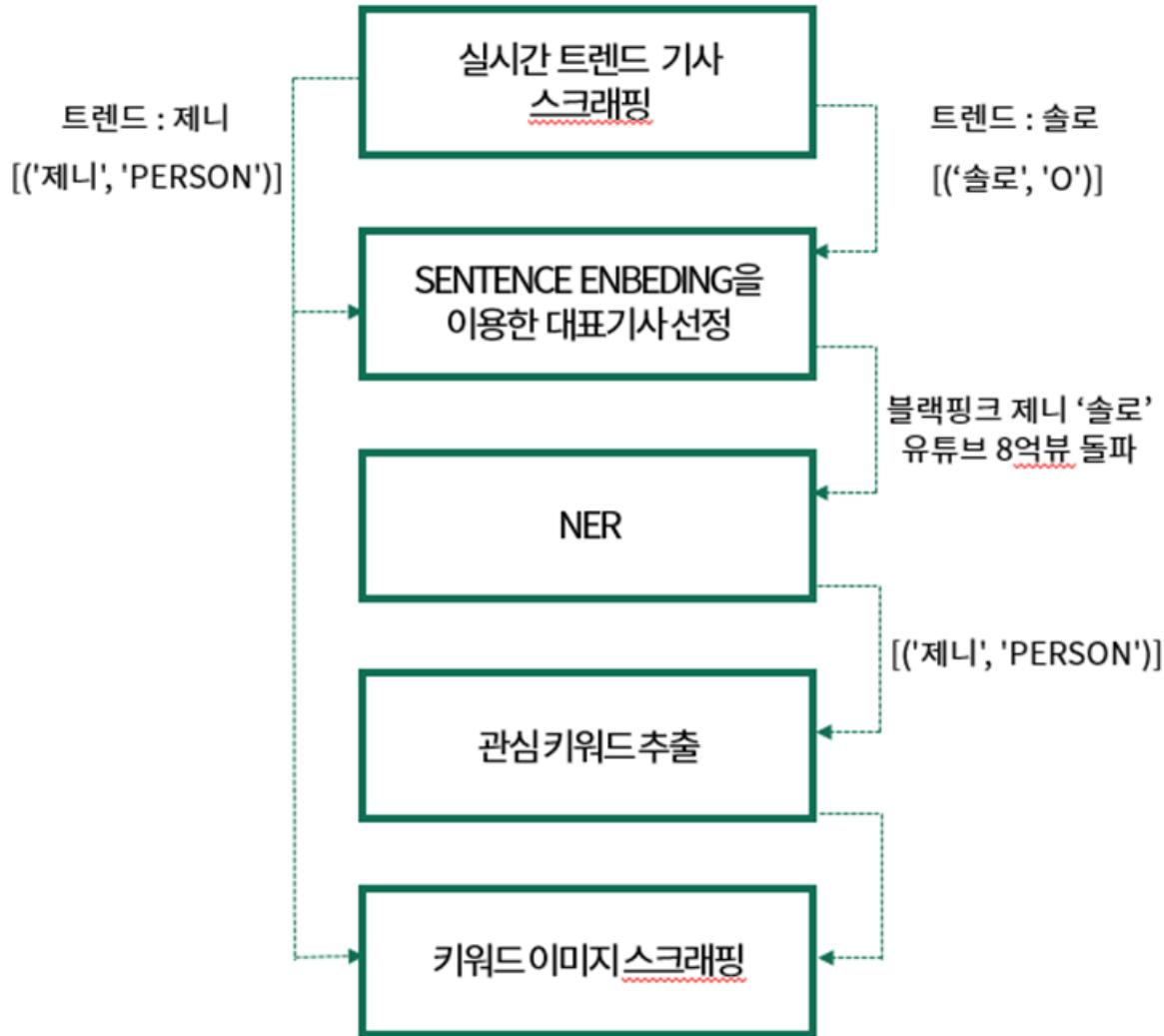
현 시간의 네이버 실시간 검색어를 제공하는 네이버 시그널의 검색어 정보를 크롤링하여 가져왔고 각 검색어에 해당하는 기사들 또한 정해진 숫자만큼 크롤링하였습니다. 기사는 특수문자를 제거한 이후 350자가 넘는 기사들을 대상으로만 크롤링 진행하였으며 검색어 당 8개의 기사를 크롤링하였습니다.

2) SENTENCE EMBEDDING - 대표기사선정

뉴스제목들 간의 Sentence Embedding을 통해 대표기사를 설정하였습니다.

8개의 기사를 가지고 왔을 때 Sentence Embedding을 각 기사의 제목에 실시하였습니다. 이 후 각기사들과 타 기사들 간의 cosine similarity를 계산하여 cosine similarity가 가장 높은 기사는 다른 기사와 유사하다고 판단했습니다. 즉 해당 검색어가 실시간 검색어에 선정된 이유와 가장 밀접한 기사라고 할 수 있는 기사라고 가정하고 해당 기사를 대표기사로 선정하였습니다.

3) NER (개체명 인식)

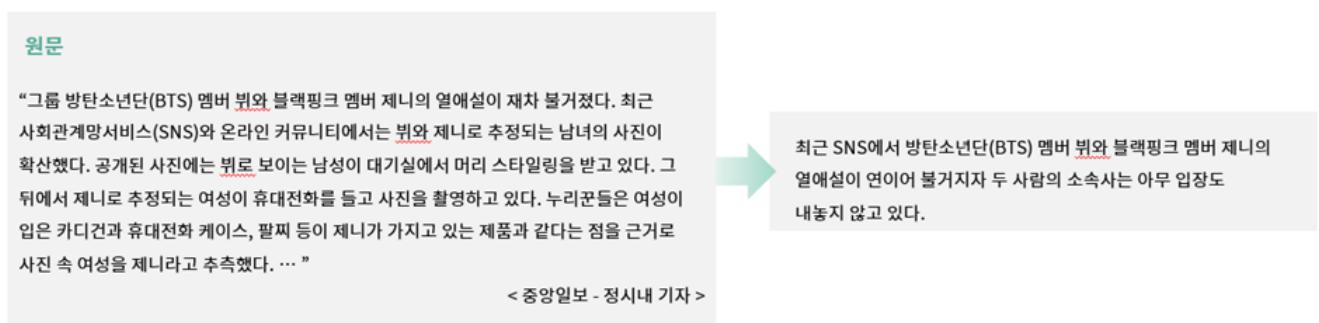


예를 들어 위의 이미지를 보면 '솔로'이라는 실시간 검색어의 대표기사의 제목에 NER을 실시하여 유의미한 이슈 키워드를 도출해냅니다.

키워드 도출에 있어서 검색어 -> 대표기사 제목 -> 대표기사 요약문(4)에서 설명)을 활용하였으며 키워드가 여러가지 도출되었을 경우에는 person -> organization -> artifact (드라마 영화 작품 등)의 우선순위로 이슈 키워드를 선정하였습니다.

4) 기사 요약문 생성

2)에서 선정한 대표기사의 본문을 대상으로 TEXT SUMMARIZATION을 실시하여 최종 결과물에 사용될 수 있도록 하였습니다. 요약은 PORORO 패키지 중 TEXT SUMMARIZATION 모델의 abstractive 방식을 사용하였고 기사 원문의 길이와 요약문의 가독성 등을 고려하여 하이퍼 파라미터 값은 beam=5, len_penalty=3, no_repeat_ngram_size=3, top_k=50, top_p=0.7, temperature=1.5로 조정하여 요약문을 추출하였습니다.



5) 대상 이미지 스크래핑

아스날



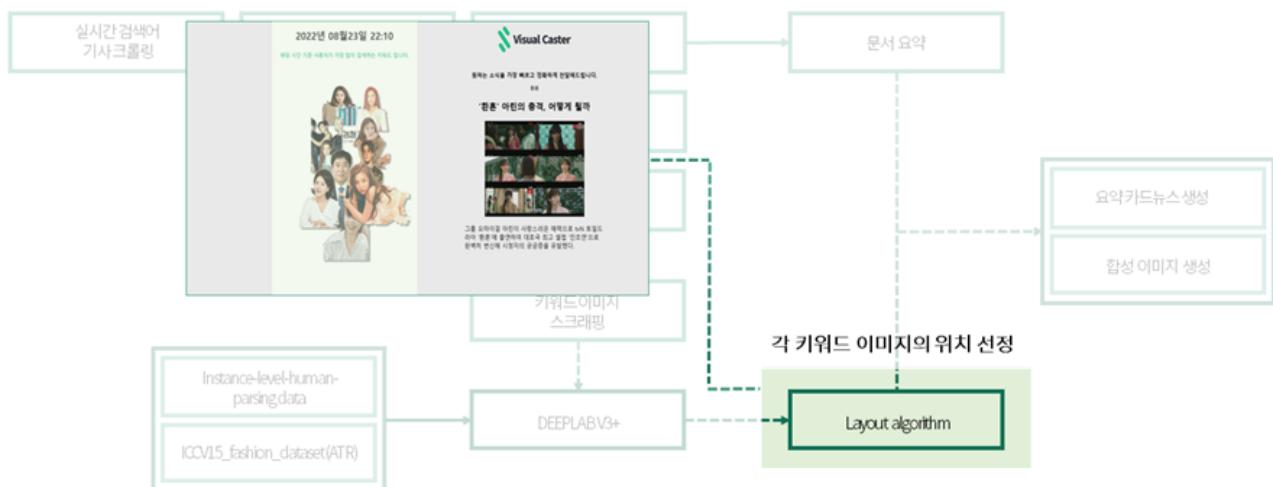
제니



나무위키 사이트 (고해상도 프로필 혹은 로고사진을 제공)에 3)에서 선정한 이슈 키워드를 검색하여 나오는 인물의 사진 혹은 단체나 작품의 사진을 스크래핑하여 이슈 키워드를 시각화 하는 데에 사용하였습니다.

8. 이미지 배치 방안

7-5) 에서 스크래핑한 이미지들은 ①단체나 지역, 창작물들의 포스터 / ②인물로 구성됩니다 ②인물의 경우 segmentation model을 통하여 하였고 ①의 경우에는 이미지 그 자체를 사용하였습니다.



1) 사이즈 조정

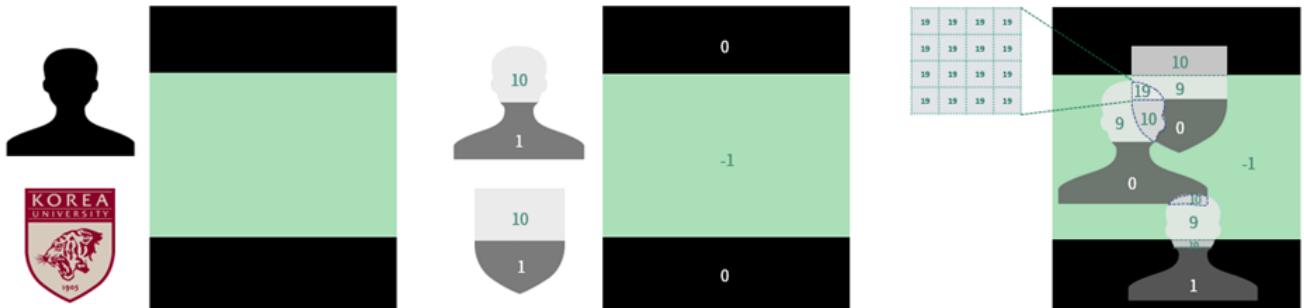
먼저 사람의 전신 이미지가 들어올 경우에는 cv2에서 제공하는 얼굴인식모델을 사용하여 얼굴부분이 상단으로 가게 crop하여 얼굴이 보일 수 있도록 하였습니다. ①의 경우 가로와 세로의 길이가 2배이상 차이 날 경우 2배이상 큰 면을 잘라내어 2배 사이즈가 되게 조정하였습니다. (ex 400x100의 이미지를 200x100의 이미지로)

segmentation된 이미지 크기가 Background size의 1/7 이 되도록 크기 조정을 해준 뒤, 검색어 순위에 맞게 한 번 더 적용하여 크기를 결정하였습니다. ①은 순위정보를 반영하지 않고 크기를 통일하였습니다.

2) 배치 가중치 생성

- 이슈 이미지의 점수

이미지를 배치하게 되면 자연스럽게 겹치는 부분이 생기게 됩니다. 이에 본 조는 이미지의 가리면 안 되는 부분(인물의 얼굴)을 설정하여 가중치를 주었습니다. 1)에서 인물의 경우 얼굴이 상단에 배치가 되게 하였기에 상단 1/2지점(Important region)은 가중치를 10으로 하단 1/2지점(Other region)은 1을 넣어주었습니다.



①의 경우 배치되는 위치에 따라 오른쪽이라면 좌측하단의 가중치가 1이되게 나머지 부분은 10이 되도록, 좌측이라면 우측상단이 1/나머지가 10이 되도록 위치에 따라 조정되도록 설정하였습니다.

- Background Score

위의 이미지처럼 상하를 기준으로 중앙부분에는 (하늘색) -1 이외의 부분은 0을 점수로 설정하였습니다.

3) 순위에 맞는 초기 위치 배치

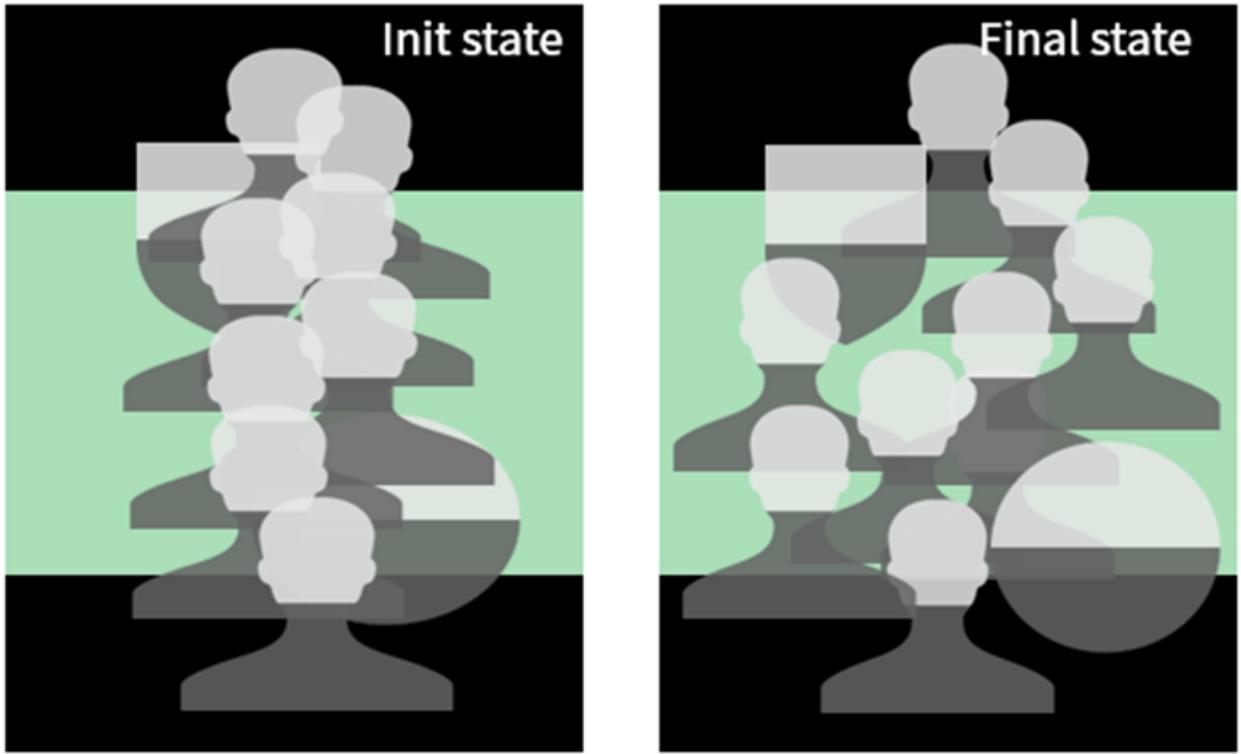
** 평가 지표

$$Arrange score(state) = \sum_{w=0}^{width} \sum_{h=0}^{height} index score(w, h)$$

$$index score = \begin{cases} -3 & pixel score = -1, 21, 22 \\ -2 & pixel score = 19, 20 \\ -10^3 & pixel score = 29, 30 \end{cases}$$

각 픽셀당 2)에서 정의한 점수를 구했습니다. 우측 상단의 이미지처럼 이미지가 겹친 경우에는 해당 위치의 점수들을 더하여 pixel score를 구했습니다.

Index score는 이미지의 중요한 부분이 많이 겹칠 경우에는 큰 페널티를 주도록 하였습니다. 위의 수식을 이용하면 이미지의 배치 상태에 대해서 arrange 스코어를 구할 수 있습니다.



arrange score를 이용한다면 중요부분이 많이 겹치지 않은 좌측 이미지가 많이 겹친 우측에 비해 더 높은 arrange score를 가질 것이고 이는 더 배치가 잘되었음을 의미합니다.

4) Neigbor와 current state 비교를 통한 Final state 반환

1위와 10위를 고정한 채로 나머지 이미지를 상하 좌우로 정해 놓은 픽셀만큼 이동했을 때를 Neigbor state라고 정의했습니다. Current state와 Neigbor state의 arrange score를 확인한 뒤 더 좋은 점수를 가지는 상태를 선택하였습니다. 이 과정을 반복하여 현재 상태의 arrange score가 이웃상태보다 높거나 threshold의 도달할 경우 Final state로 반환하였습니다.

9. 최종 결과물

7-4)와 8의 이미지 레이아웃 좌표를 받아 django를 이용하여 좌표, segmentation이후 이미지, 요약문 등을 담은 csv 형태를 웹으로 전송하여 웹으로 구현하였습니다.

Visual Caster
2022년 8월 23일 화요일 오후 10:10

1 김해경 의혹 출석
3 환흔
4 김대기 사의 표명한
5 산이 결혼
6 활금기면
7 아이브
8 총기인 청주콘서트
9 검찰 쟁방을 전
10 한혜진

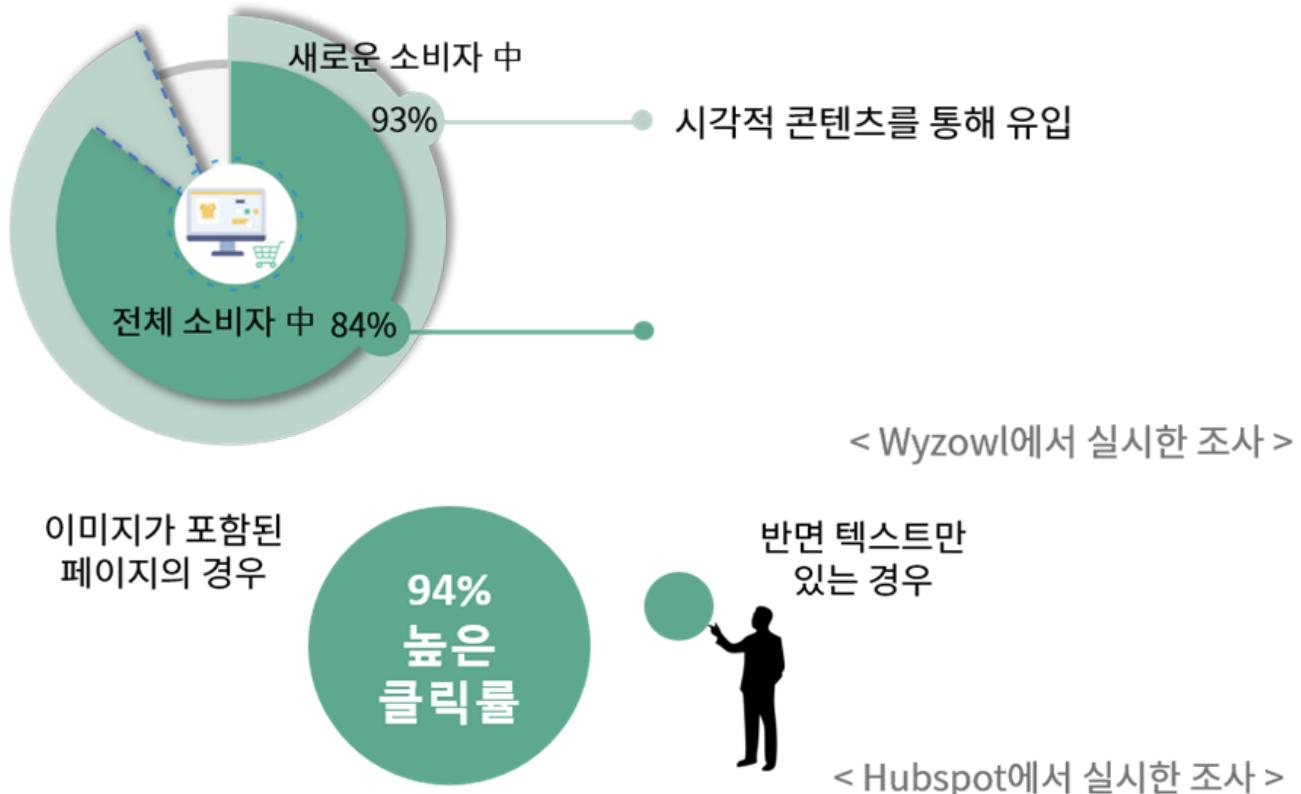
이미지로 보고싶진 않으신가요?



첫 화면에서 제공되는 키워드 클릭 시 '이미지로 보고싶진 않으신가요?'라는 문구를 띠웠고 이를 릭하게 되면 실시간 검색어가 시각화된 페이지로 이동하게 됩니다. 해당 페이지에서 이미지로 나타낸 검색어에 마우스를 올리게 되면 작은 팝업으로 순위와 원래의 검색어가 나오게 됩니다. 우측에는 해당 검색어가 현재 이슈가 된 이유를 잘 설명하는 뉴스를 카드 뉴스 형식으로 나타내어주고 이미지 클릭 시 기사의 원본 링크로 이동하게 됩니다.

10. 서비스 활용 방안 및 기대효과

1) 상업적 효과 극대화 가능



미국 Finances Online에서 진행한 다양한 소셜 미디어(Buffer App, BuzzSumo, Medium, NewsCred, Forbes)의 통계 수치 기반 조사에 따르면, 이미지 및 영상을 포함한 컨텐츠의 경우 단순 텍스트 컨텐츠에 비해 훨씬 높은 조회수 및 인용 수를 기록한 것을 볼 수 있었습니다. 또한, Wyzowl에서 실시한 조사에 따르면 84%의 소비자가 시각적 프로모션 컨텐츠를 통해 구매를 결정하게 되었다고 대답하였고, 새로운 소비자 중 93%가 시각적 컨텐츠를 통해 유입되었다고 대답하였습니다. 이런 결과들을 근거로, 현대 마케팅 시장에서는 이미지와 글을 통합하여 사용하는 것이 이용자들의 관심을 끄는데 효과적이며 나아가 실제 판매 및 더 나은 소비자 경험을 제공하는데 중요하다고 판단합니다. 또한 Hubspot에서 실시한 조사에 따르면 이미지가 포함된 페이지의 경우, 텍스트만 있는 경우보다 94% 클릭률이 높았다고 합니다.

저희 서비스는 위와 같은 마케팅 전략에 부합합니다. 소비자에게 중요한 이미지만 오려내어 시각적 정보를 전달함과 동시에 부가적인 글자 정보를 추가 제공함으로써, 소비자들의 관심을 끄는 동시에 명확한 정보전달과 관심을 끄는데 효과적이며 판매량 증가 및 소비자들의 만족을 성공적으로 이끌어 낼 수 있습니다.

2) 실시간 트렌드 제공

실시간으로 바뀌는 스포츠 리그, 음원 사이트, 배달어플 등 순위 정보가 제공되는 서비스에서의 실시간 트렌드 적용은 좋은 효과를 가져올 수 있습니다.



위의 이미지는 스포츠 리그에서의 실시간 트렌드 적용입니다. 스포츠 경기에서는 해당 경기의 순위대로 선수들을 보여줄 것입니다. 카드뉴스에는 선수의 경력이나 특징을 제공합니다. 프리미어 리그를 예로 들면 10위권 내의 구단 대표 소속 선수 이미지를 보여줄 수 있습니다. 카드뉴스에는 해당 구단의 승점, 경기 일정 등 정보를 제공됩니다.

3) 위치 기반 관광지 추천



우리 나라의 관광지나 문화재의 경우, 외국인 관광객은 명칭만으로 어떤 장소인지 쉽게 정보를 알 수 없습니다. 이는 외국인 관광객을 타겟팅으로 한 서비스로 지도 상의 이미지를 통해 추가적인 정보를 알 수 있도록 하고 이미지와 함께 같이 제공되는 카드 뉴스를 통해 현재 위치로부터의 거리 등 추가적인 정보를 제공해줍니다. 이 서비스는 언어가 아닌 이미지를 통해 정보에 대한 접근을 쉽게 할 수 있으며 언어적 한계를 극복할 수 있습니다.

4) 여행 가능 국가 정보 제공



팬데믹 상황이 장기간 지속되면서 해외여행을 원하는 사람들이 늘고 있지만 아직까지는 해외 출입국이 자유롭지 않고 국가마다 방역 수칙과 출입국 심사 기준이 계속해서 달라지고 있습니다. 따라서 세계지도에 입국 가능, 불가능 국가를 색으로 구분하여 표시하고, 사용자가 현재 입국 가능한 나라에 마우스 오버하면 입국 시 필요한 절차 정보를 카드뉴스로 제공합니다. 또, 해당 나라의 명소 이미지를 가져와 시각화하여 제공합니다. 이와 같은 복합정보를 시각화하여 실시간 현황 파악에 용이합니다.

이와 같이 실시간 트렌드 서비스는 여러 분야에서의 적용이 가능하며 다양한 기대 효과를 가져올 수 있습니다.