

2020년 인공지능 학습용 데이터 교육영상 “K-Fashion 이미지” 소개 및 활용

2021.03

주식회사 웨얼리
이사 김사현





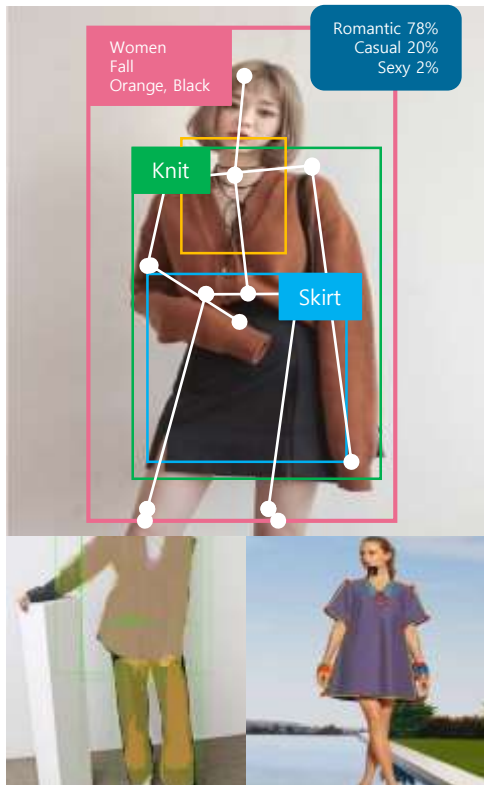
INDEX

01. 인공지능 학습용 데이터 소개
02. 저작도구 활용법
03. 데이터 학습 방법 및 시연
04. 학습된 모델 결과 확인 방법 및 시연
05. 서비스 개발 시 학습된 모델 활용 방법

01 인공지능 학습용 데이터 소개

데이터 소개

K-Fashion 학습용 이미지 데이터 구축 사업을 통해 패션 트렌드에 민감한 20~30대 여성 패션 이미지 데이터를 확보하여 인공지능 학습용 데이터셋을 약 120만 건 구축하였음



데이터 구축 방법 및 목표

- 데이터 구매를 통하여 어노테이션을 위한 이미지 데이터 확보
- 시맨틱 세그멘테이션, 일반/전문 패션요소가 입력된 120만장의 학습용 데이터셋 구축 (대분류 10가지, 세부속성 186개, 스타일 23가지)

데이터 활용 방법

- 인공지능영역을 활용한 패션 이미지 시각지능 기술 개발
- 패션 아이템과 스타일 분류 API를 활용한 이미지 자동 태깅 및 검색 서비스
- 패션 아이템 추천 API 활용한 패션 스타일 추천 서비스

01 인공지능 학습용 데이터 소개

데이터 가공방법

인공지능 학습용 데이터는 데이터 수집, 레이블링, 활용 등 3단계의 프로세스로 구축·가공됨
 특히, 레이블링은 일반인(1차), 전문가(2차)로 진행하여 산·학계 모두 활용가능한 항목으로 구성됨

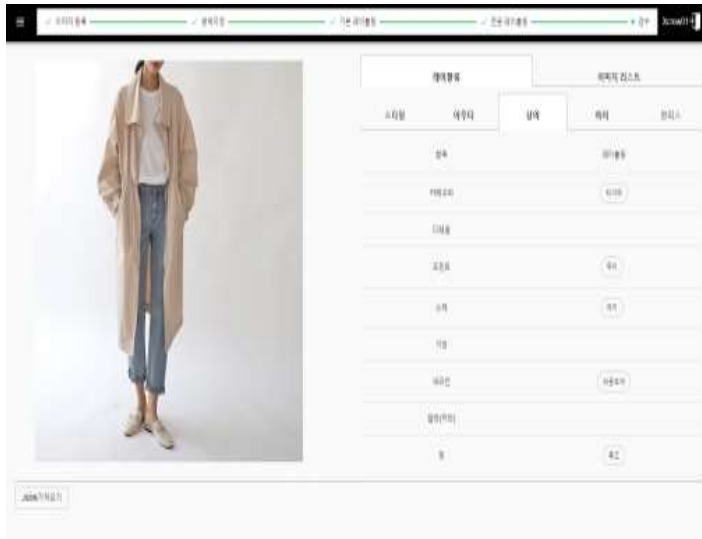


02 저작도구 활용법

데이터 구축방법

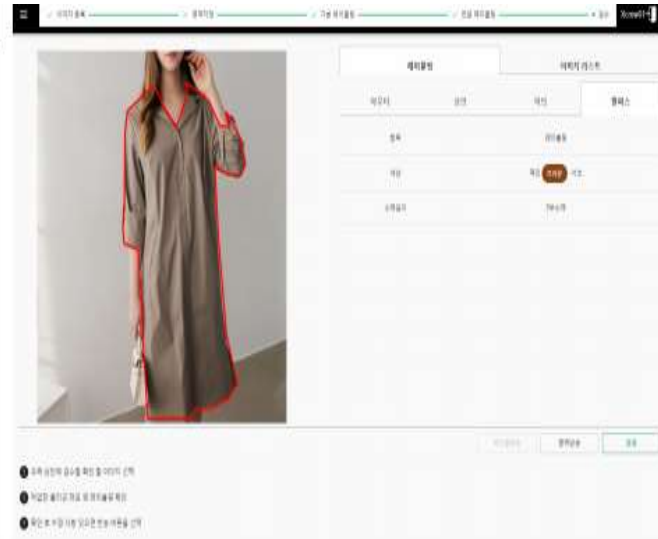
이미지 데이터를 수집이후 학습용 이미지 데이터 요건에 적합한 이미지를 선별함. 그 후 저작도구 서비스에 이미지들은 연결하여 일반/전문 레이블링을 진행한 후 최종 관리자가 데이터를 검수함

전문가 레이블링
(의류학과 학사, 석사, 박사 등)



스타일, 디자인, 소재 등 전문 레이블링

일반인 레이블링
(클라우드소싱)



영역지정 및 색상, 길이 등 기본 레이블링

01 인공지능 학습용 데이터 소개

데이터 제공형식

데이터는 이미지와 레이블링 파일(json형식) 쌍으로 제공하며, json파일은 정해진 룰(rule)에 따라 레이블링 결과를 정리하여 제공됨

패션 AI 데이터 데이터 형식



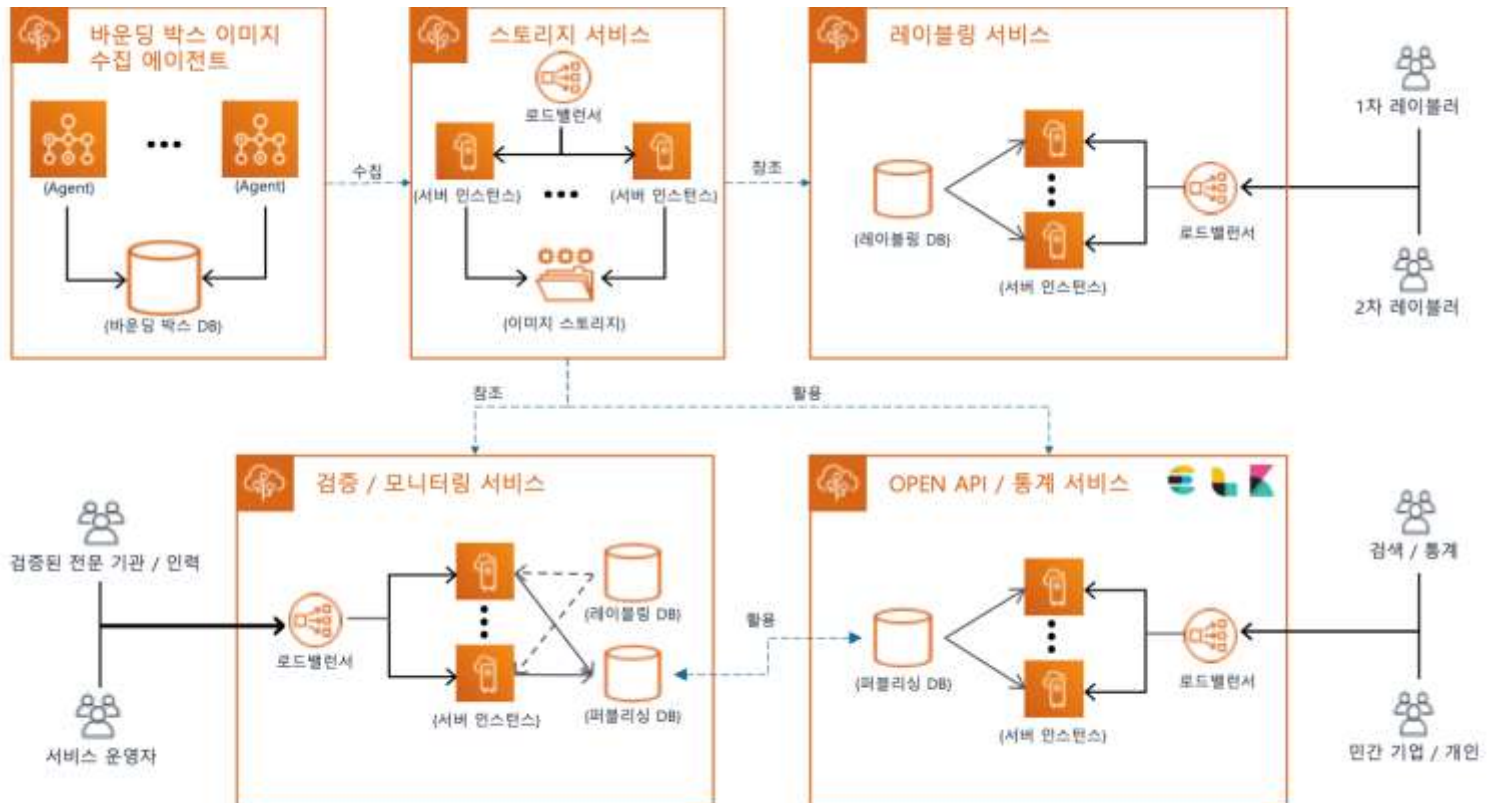
```
{
  "이미지 정보":{
    "이미지 식별자":16566,
    "이미지 높이":1066,
    "이미지 파일명":"u_14859239888517400_1493917978.jpg",
    "이미지 너비":800
  },
  "데이터셋 정보":{
    "파일 생성일자":"2020-09-18 17:07:48",
    "데이터셋 상세설명":{
      "렉트좌표":{
        "아웃터":{()
      },
      "하의":{["X좌표":210.5,"Y좌표":329.5,"가로":322,"세로":375],
        "원피스":{()
      },
      "상의":{["X좌표":180.5,"Y좌표":15.5,"가로":325,"세로":370]},
      "폴리곤좌표":{["아웃터":{()
      },
      "하의":{["X좌표1":222.0,"Y좌표1":522.0,..."X좌표5":256.0,"Y좌표5":705.0]},
        "원피스":{()
      },
      "상의":{["X좌표2":273.0,"Y좌표4":69.0,..."X좌표38":385.0,"Y좌표38":506.0]},
      "라벨링":{["스타일":{["스타일":"리조트","서브스타일":"컨트리"}]
      },
      "아웃터":{()
      },
      "하의":{["기장":"발목","색상":"블랙","카테고리":"팬츠","디테일":"리본"],
        "소재":["울/캐시미어"],"프린트":["아가일"],"핏":"루즈"},
        "원피스":{()
      },
      "상의":{["기장":"롱","색상":"블루","카테고리":"브라탑","옷깃":"차이나칼라",
        "디테일":["리본"],"소매기장":"긴팔","소재":["실크"],
        "프린트":["하트"],"넥라인":"브이넥","핏":"루즈"}]
      },
      "파일 번호":16566,
      "파일 이름":"u_14859239888517400_1493917978.jpg"
    }
  }
}
```

02 저작도구 활용법

저작도구 설명

K-Fashion 학습용 이미지 데이터 구축에 사용된 저작도구는 다양한 이미지 저작도구를 검토하고 장점을 취합하여 K-Fashion 맞춤형 저작도구로 컨소시엄에서 직접 개발하였음

저작도구 구성 개요도



02 저작도구 활용법

저작도구 활용법

K-Fashion 저작도구는 일반인이 작성할 수 있는 일반 레이블링(세그멘테이션, 일반라벨링) 작업과 전문가가 작성할 수 있는 전문 레이블링으로 구분하여 작업이 이루어질 수 있도록 함
관리자 및 검수기능을 활용하여 작업결과의 품질과 작업량을 관리할 수 있음

1 일반레이블링 - 세그멘테이션



2 일반레이블링 - 일반 라벨링



3 전문가레이블링 - 전문 레이블링



4 작업결과 확인



5 작업자 및 작업량 관리



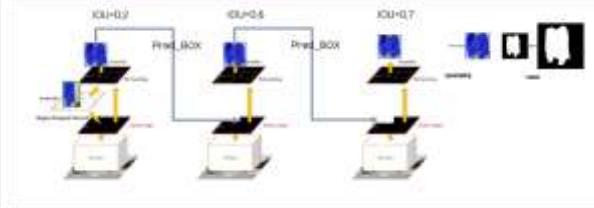
03 데이터 학습 방법 및 시연

패션 아이템 검출 모델 이미지 내 패션 아이템과 관련된 시각적 요소를 학습한 후, 사용자가 새로운 이미지를 쿼리로 사용할 경우 다양한 패션 이미지에서 패션 아이템의 위치를 탐지하고 아이템 모양을 구하는 검출 모델을 개발함

데이터 준비 및 전처리



학습 알고리즘



Cascade R-CNN 모델

데이터 학습 진행



학습 결과 확인



03 데이터 학습 방법 및 시연

패션 아이템 속성 분류 모델

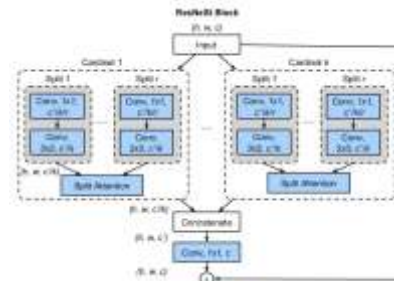
패션 아이템 특성상 하나의 아이템에 대해 여러 속성이 존재하며 이를 분류하기 위해 많은 인력이 소모됨. 이를 해결하기 위한 모델로 패션 아이템 속성 분류 모델을 개발 및 구현함

데이터 준비 및 전처리



```
# Preprocessing
python main/convert_images_to_json.py --image_paths './data/images/fashion/' \ # 경로
--photo='category' \ # 추출하고자 하는 속성 정보를 준다
--output_paths './data/json' \ # 이미지 및 속성 정보를 저장할 위치를 준다
```

학습 알고리즘



ResNet의 residual block

데이터 학습 진행

```
python main/train.py --data='./data/json/category' \ # 각 속성 분류 모델 인풋 위치
--n_classes=21 \ # 학습하는 속성의 레이블 개수
--epochs=100 \ # 최대 epoch 범위
--batch-size=64 \ # batch size 범위
--workers=12 \ # worker 개수 범위
--learning-rate=0.001 \ # learning rate 범위
--save_model_path './checkpoint/' \ # 학습된 모델을 저장할 위치
```

```
# 실행 로그 예시
../checkpoint/category
vs so checkpoint found at ''
lr: [0.001]
Training: 17% [██████████] | 8/35 [00:13<01:00, 2.23s/it]
```

학습 결과 확인

```
python main/test.py --data='./data/json/category' \ # 테스트 인풋 위치
--evaluate = True \ # evaluate 테스트 모드
--save_model_path './checkpoint/category/model_category_test.pth.tar' \ # 불러올 모델 위치

# 실행 로그 예시
WARNING:
<- loading checkpoint './checkpoint/category/model_category_test.pth.tar'
Test: 100% [██████████] 21/21 [00:00<00:00, 2.00s/it]
Top 3 accuracy: 0.734
[100%] category class speed: 100.0720
```

03 데이터 학습 방법 및 시연

스타일 분류 모델

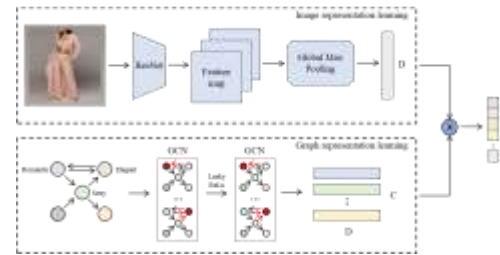
패션 이미지는 여러 아이템 매치에 따라 다양한 스타일이 존재하여 정확한 스타일 분류를 위해 다중 스타일 라벨(Label)이 고려된 알고리즘을 활용하여 분류 모델을 개발함

데이터 준비 및 전처리



```
# Preprocessing
python data/environment_model_train.py --image_path './data/images/fashion/' --c 클래스
--image_style './data/images/' --c 후속이코드가 있는 파일 정보를 의미
--output_path './data/style/' --c 이미지 등 속성 정보를 저장할 위치를 의미
```

학습 알고리즘



ML-GCN구조를 적용한 아키텍처

데이터 학습 진행

```
python main/train.py --data './data/json/style/' --c 이미지 정보 위치
--image './data/graph/custom_sovden_final.pkl' --c 워드백인 연결 위치
--adj './data/graph/custom_adj_final.pkl' --c Adjacency matrix 위치
--n_classes 21 --c 학습하는 속성의 레이블 개수
--epochs 100 --c 최대 epoch 값
--batch_size 64 --c batch size 정보
--workers 12 --c worker 개수 정보
--learning_rate 0.001 --c learning rate 정보
--learning_rate_scheduler = 0.01 --c 사전에 학습된 레이어에 대한 learning rate 정보
--save_model_path './checkpoint/fashion_style/' --c 학습된 모델을 저장할 위치

# 실행 로그 예시
.../checkpoint/fashion_style/
python checkpoint_train.py
Training: 200% [████████████████████] 379/379 [00:09:00.00, 1.54it/s]
Test: 100% [████████████████████] 01/02 [00:12:00.00, 2.46it/s]
Test: Loss 8.6327 MAP 25.722
DP: 0.1867 IR: 0.9996 IP1: 0.1295 IP: 0.1860 CR: 0.8993 FP: 0.5147
OP_1: 8.3289 OR_1: 0.7791 OP1_1: 0.6072 OP_2: 0.2718 OR_2: 0.6524 FP1_1: 0.3900
save model checkpoint/fashion_style/checkpoint.pth.tar
```

학습 결과 확인

```
python main/test.py --data './data/json/style/' --c 테스트 데이터 위치
--image_style './data/' --c image 테스트 정보
--save_model_path './checkpoint/fashion_style/model_test.pth.tar' --c 불러올 모델 위치

# 실행 로그 예시
Starting...
python checkpoint_train.py --checkpoint ./checkpoint/fashion_style/model_test.pth.tar
Test: num [████████████████████] 02/02 [00:00:00.00, 1.00it/s]
Top-5 recall: 0.5011
Top-5 email: 0.8888
[Done] Total time spent: 103.1036
```

03 데이터 학습 방법 및 시연

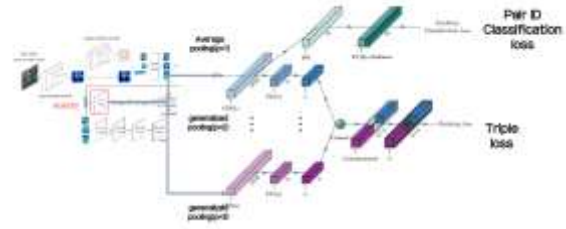
패션 아이템 추천 모델

패션 이미지 내에 존재하는 다양한 패션 아이템의 정보를 추출하고 추출된 아이템과 어울리는 다른 범주의 카테고리의 아이템을 추출하여 제공하는 학습모델임

데이터 준비 및 전처리



학습 알고리즘



Combined Global Descriptors 모델 구조

데이터 학습 진행

```

python3 train.py --Fashion_recommend --data_dir /home/kore
2020-11-25 09:58:35.915 - Loading abstract from /home/kore
2020-11-25 09:58:35.944 - Extracting fea /dataset/women/00991.jpg
2020-11-25 09:58:36.007 - Extracted 3 items: ["T-shirts"]
2020-11-25 09:58:36.094 - Load egl model from /model/egl_model.pt
Query time type: 1.011715, 4 items recommended
2020-11-25 09:58:36.092 - Saved json to /dataset/res_images/2020_11/0991.json
    
```



```

{ "resultCode": "SUCCESS",
  "resultMessage": "SUCCESS",
  "body": { "filePath": "/home/kore
a/fashion-recommendation/dataset
/seg_images/images/056665.jpg" },
  "responseDate": "2020-11-25 14:
35:19"}
    
```

학습 결과 확인



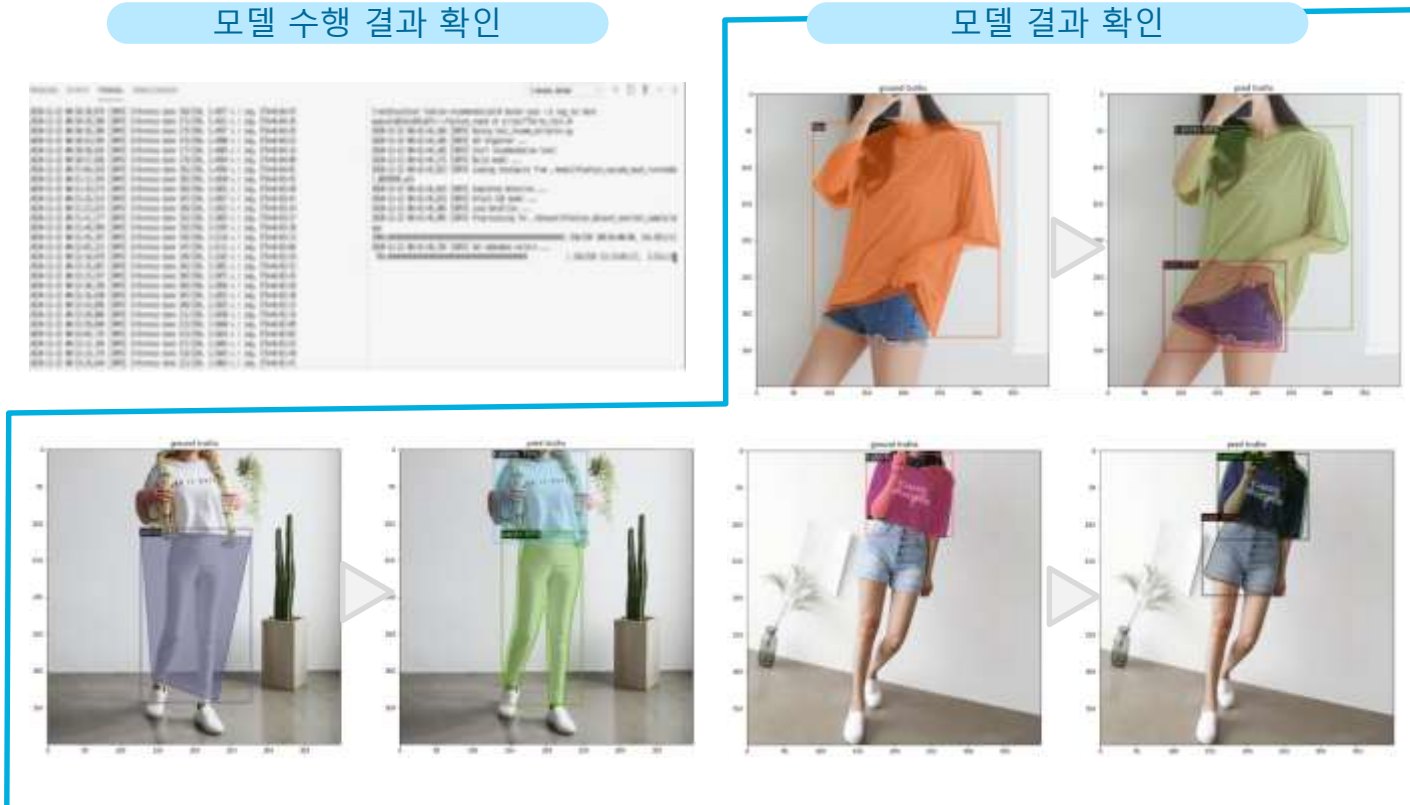
04 학습된 모델 결과 확인 방법 및 시연

패션 아이템 검출 모델 패션 아이템 검출 모델을 통해 나타나는 결과는 패션 아이템 영역을 확인하고 해당 아이템이 무엇인지 그림과 검출된 내용에 대한 정보를 담은 파일(json)으로 구성되어 제공됨

모델 수행 결과 확인



모델 결과 확인



04 학습된 모델 결과 확인 방법 및 시연

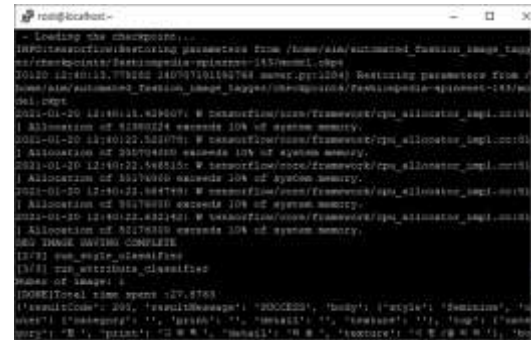
패션 아이템 속성
분류 모델

이미지 내에 패션 아이템에 대해 다양한 특성(성별, 계절감, 형태, 색상 등)이 존재하여 하나의 아이템에 대하여 자동으로 특성을 분류하기 위한 모델로 이미지와 쌍을 이루는 산출물(json)이 구현

모델 수행 결과 확인



모델 결과 확인



```

{
  "resultCode": 200,
  "resultMessage": "SUCCESS",
  "body": {
    "style": "feminine",
    "outer": {
      "category": "타입",
      "print": "그래픽",
      "detail": "리본",
      "texture": "시퀀/글리터"
    },
    "top": {
      "category": "셔링",
      "print": "시퀀/글리터",
      "detail": "셔링",
      "texture": "시퀀/글리터"
    },
    "bottom": {
      "category": "스커트",
      "print": "셔링",
      "detail": "셔링",
      "texture": "시퀀/글리터"
    },
    "onepiece": {
      "category": "셔링",
      "print": "셔링",
      "detail": "셔링",
      "texture": "시퀀/글리터"
    },
    "responseDate": "2021-01-21 02:40:02"
  }
}
    
```


04 학습된 모델 결과 확인 방법 및 시연

패션 아이템 추천 모델

패션 아이템 추천 모델은 입력된 이미지와 유사한 스타일의 아이템을 찾은 후 사용자가 선택한 속성 카테고리(색상, 길이, 소재 등)를 적용하여 가장 유사한 이미지를 찾아 제공함

모델 수행 결과 확인



모델 결과 확인



상의 : 셔츠, 셔츠칼라, 루즈, 체크
하의 : 스커트, 미니, 데님



변경

미니 → 롱

상의 : 셔츠, 셔츠칼라, 루즈, 체크
하의 : 스커트, 롱, 데님



변경

셔츠 → 티셔츠

상의 : 티셔츠, 셔츠칼라, 루즈, 체크
하의 : 스커트, 롱, 데님



05 서비스 개발 시 학습된 모델 활용 방법

패션 이미지 자동 태깅 서비스

패션 아이템 특성 분류 모델을 활용하여 입력 받은 패션 이미지에서 패션 아이템 탐지 후 이에 대한 아이템에 대한 속성분류를 통해 자동으로 이미지에 태깅하는 서비스임

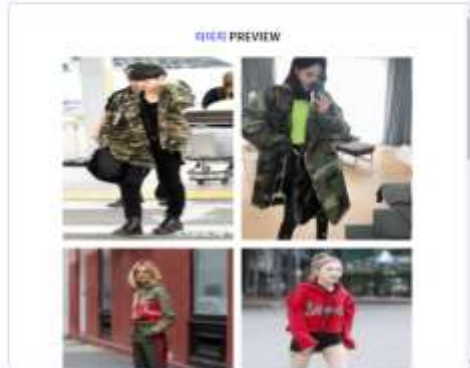
Tagging Service

패션 이미지 또는 폴더 링크를 입력하여 해당 패션 아이템의 분석결과를 제공합니다.



Analysis Result

업로드된 분석결과를 원하는 해상도로 다운로드하세요.



조회 결과

총 2건의 이미지 태그완료

REFRESH COPY LINKS DELETE

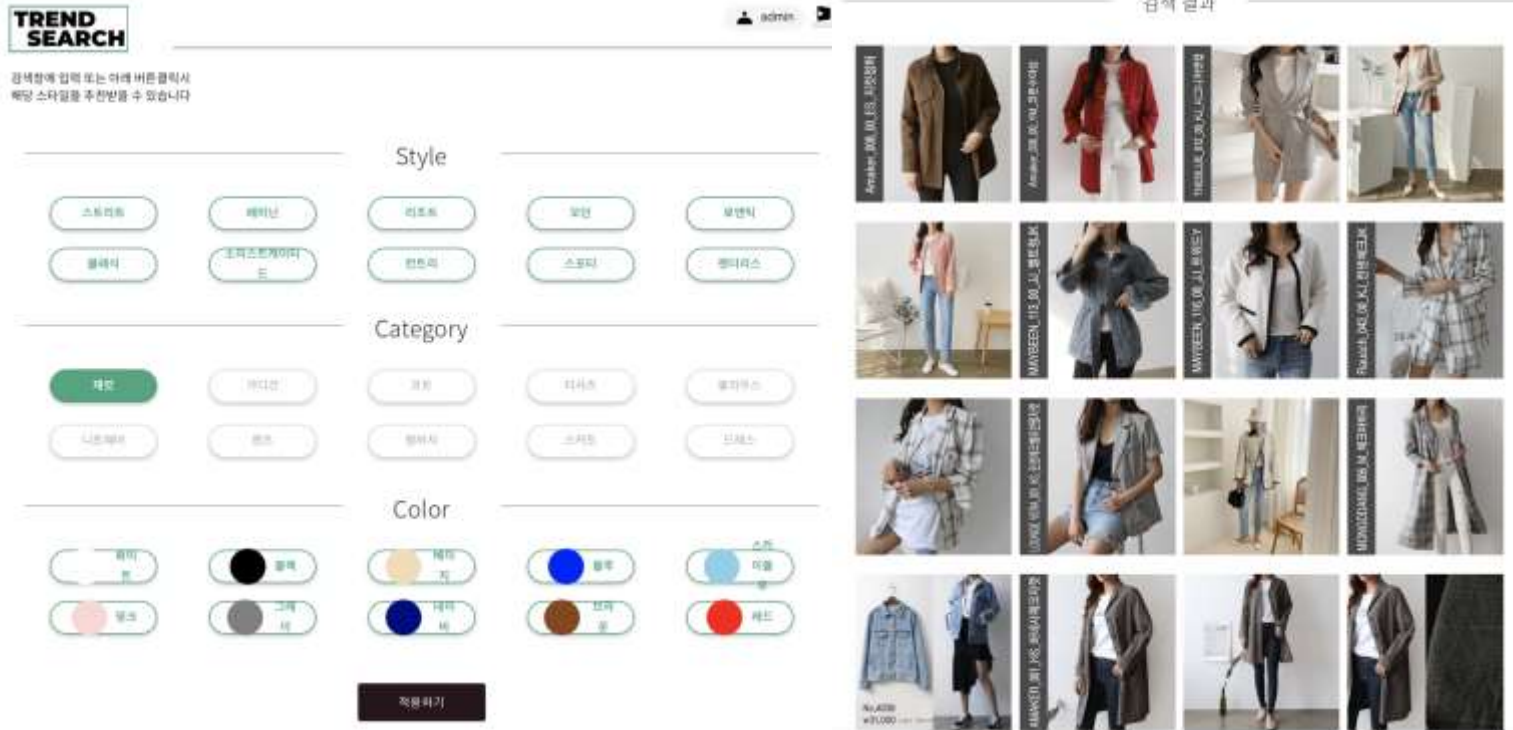
파일명	스타일	카테고리 (여우)	프린트 (여우)	다재업 (여우)	소재 (여우)	카테고리 (남자)	프린트 (남자)	다재업 (남자)	소재 (남자)	카테고리 (키)	프린트 (키)	다재업 (키)	소재 (키)
01.png	casual	패션	여가용	비데업	시폰/양모	베스트	여가용	서킷	양면	팬츠	그림자	서킷	시폰/양모
02.png	casual	오토	양자	리본	양면	베스트	양자	서킷	양면	스커트	양자	비데업	가죽

접속주소 - <http://kfashion.ai/tagging>

05 서비스 개발 시 학습된 모델 활용 방법

트렌드 모니터링을 위한 패션 이미지 검색 서비스

패션 이미지 검색 서비스는 사용자가 입력한 텍스트 혹은 이미지 정보를 입력 받고, 기 학습된 아이템 탐지모델과 속성 분류를 통해 입력된 태그에 해당하는 유사한 이미지 제공하는 서비스 임




.....05 서비스 개발 시 학습된 모델 활용 방법.....

패션 아이템 추천 모델

매칭 아이템 추천 서비스는 사용자의 선호도와 제공한 이미지를 입력 받아, 요청한 패션 아이템에 대해 학습용 이미지 데이터 중 가장 유사한 이미지 제공하는 서비스임

image upload

매칭하고 싶은 이미지를 업로드해주세요



업로드할 이미지 (max)를 클릭해주세요!

어떤 분과 매칭할까요?

매칭할 대상별

여성	가디언	연희	진숙희	황영희
김민정	연우	정혜리	조기림	김

스타일

트렌디	캐주얼	직접과	고급	방한복
스포츠	캐주얼	정제과	스포츠	가디언
정제과	스포츠	정제과	스포츠	가디언


컬러

									
---	---	--	--	---	--	--	--	---	---

꼭 눌러주세요!

*** 데이터의 최종값은 30일마다 갱신됩니다.

내가 올린 이미지



검색 결과



감사합니다

주식회사 웨얼리
이사 김사헌

E-MAIL :
simon@wearly.co.kr
PHONE : 010-9165-4141

