

문서명 01_README.pdf		작성일 2022.07.05.	전체 페이지수 7
소 속 시각지능연구실	작성자 박민호	제 목 챌린지 서브태스크 1 베이스라인 설명	문서관리자 송화전

챌린지 서브태스크 1

베이스라인 설명

CONTENTS

1. 문서의 목적	3
2. DB.....	3
2.1 패션 아이템의 이미지 DB.....	3
2.2 패션 아이템의 라벨 정보 DB.....	4
3. 파일 구조.....	4
4. 실행	5
4.1 실행 환경	5
4.2 설치된 라이브러리	5
4.3 실행 방법	5
4.4 파라미터 옵션	5
5. 베이스라인 설명	6
5.1 개요	6
5.2 학습 및 평가.....	7
6. 라이선스.....	7
7. 연락처.....	7

1. 문서의 목적

- 본 문서는 “2022 ETRI 자율성장 인공지능 경진대회”에서 서브태스크 1 을 위한 베이스라인을 설명하는 문서이다.

2. DB

2.1 패션 아이템의 이미지 DB

- 패션 아이템의 이미지 데이터들은 각각 다른 사이즈의 이미지들로 구성되어 있고, 이미지 예시는 (그림 1)과 같음.

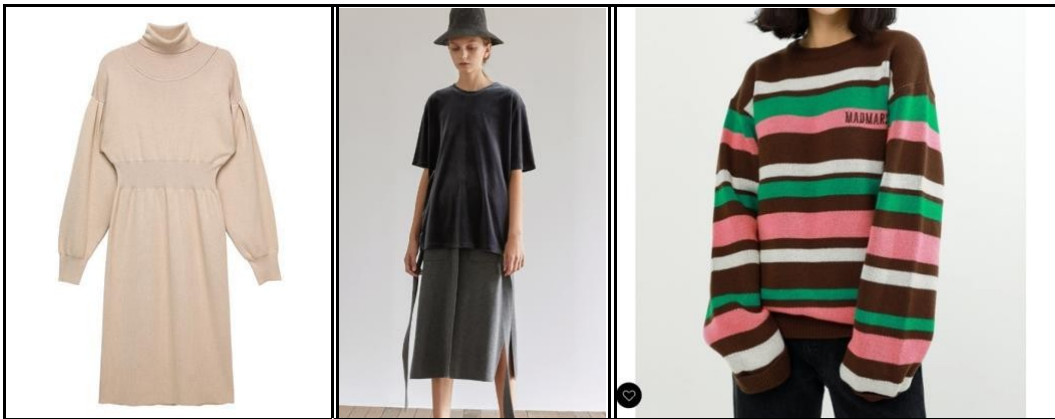


그림 1. 패션 이미지 예시

- 각 패션 이미지들은 해당 종류에 맞게 분류되어 폴더에 저장되어 있고, 그 형태는 (그림 2)와 같음.

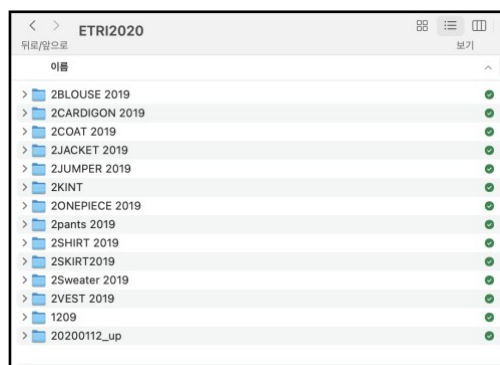


그림 2. 이미지 DB 폴더 구성

2.2 패션 아이템의 라벨 정보 DB

- 각 패션 이미지들은 그에 해당하는 라벨 정보를 가지고 있음. 라벨 종류에는 의상 특징 중 하나인 감성특징들로 일상성, 성, 장식성으로 구성됨.
- 의상 이미지들은 이미지 내에 의상이 위치하고 있는 부분을 가리키는 정보가 필요함. 이에 대한 정보로 바운딩 박스(bounding box) 좌표를 데이터 라벨에 포함하였음. (그림 3)

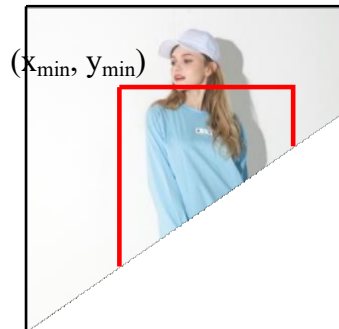


그림 3. 바운딩 박스를 포함한 이미지 예시

- 라벨 정보 DB 예시는 (그림 4)와 같음.

	A	B	C	D	E	F	G	H
1	image_name	BBox_xmin	BBox_ymin	BBox_xmax	BBox_ymax	Daily	Gender	Embellishment
2	1209/BC000001-1.jpg	45	197	313	393	4	5	0
3	1209/BC000002-1.jpg	117	114	476	556	1	5	1
4	1209/BC000003-1.jpg	62	212	372	544	1	5	0
5	1209/BC000004-1.jpg	159	17	484	395	1	5	1
6	1209/BC000005-1.jpg	51	54	561	441	1	3	0
7	1209/BC000006-1.jpg	2	38	678	870	1	3	0
8	1209/BC000007-1.jpg	167	184	435	444	4	5	0
9	1209/BC000008-1.jpg	116	205	404	425	1	3	0
10	1209/BC000009-1.jpg	124	239	380	423	1	3	0
11	1209/BC000010-1.jpg	22	104	499	418	4	5	0
12	1209/BC000011-1.jpg	189	231	357	463	4	5	0
13	1209/BC000012-1.jpg	106	240	426	641	3	4	0
14	1209/BC000013-1.jpg	118	171	402	583	4	5	0
15	1209/BC000014-1.jpg	30	59	611	777	1	5	0
16	1209/BC000015-1.jpg	93	193	446	614	4	5	0
17	1209/BC000016-1.jpg	46	75	574	654	4	5	0
18	1209/BC000017-1.jpg	125	144	404	405	1	5	0
19	1209/BC000019-1.jpg	59	67	451	648	1	3	0
20	1209/BC000020-1.jpg	14	6	702	725	1	5	0
21	1209/BC000021-1.jpg	107	40	395	435	1	5	0
22	1209/BC000022-1.jpg	79	192	427	656	1	5	0
23	1209/BC000023-1.jpg	84	74	361	386	1	5	0
24	1209/BC000024-1.jpg	57	57	385	617	1	5	0
25	1209/BC000025-1.jpg	103	154	331	474	1	5	0
26	1209/BC000026-1.jpg	55	82	336	503	1	5	1
27	1209/BC000027-1.jpg	127	108	453	561	1	5	0
28	1209/BC000028-1.jpg	59	180	629	480	1	5	0
29	1209/BC000029-1.jpg	23	120	413	530	1	1	0
30	1209/BC000030-1.jpg	37	13	518	828	3	1	0
31	1209/BC000033-1.jpg	146	0	498	499	4	5	2
32	1209/BC000034-1.jpg	0	91	712	727	1	5	1
33	1209/BC000035-1.jpg	82	60	569	407	1	5	0

그림 4. 라벨 정보 DB 예시

3. 파일 구조

- ./dataset.py
패션 아이템 이미지 DB 및 라벨 정보 DB를 불러오는 소스코드

- `./network.py`
패션 아이템 이미지의 속성 분류를 수행하는 모델 구조를 정의하는 소스코드
- `./test.py`
학습된 속성 분류 모델을 사용하여 원하는 패션 이미지의 속성을 추출하고 모델 성능을 평가하는 소스코드
- `./train.py`
패션 이미지 속성 분류 모델을 사용자가 설정한 파라미터대로 학습을 진행하는 소스코드

4. 실행

4.1 실행 환경

Ubuntu 18.04
Python3.6
CUDA 10.2
CUDNN 8.0.4
PyTorch 1.7.0

4.2 설치된 라이브러리

Torchvision 0.8.1
Scikit-image 0.17.2
Scikit-learn 0.24.1
Pandas 1.1.5

4.3 실행 방법

- 패션 이미지 DB 를 활용하여 학습을 하고자 할 때
train.py 소스코드 내에 원하는 파라미터 설정 후 실행
- 학습된 모델을 검증하고자 할 때
test.py 소스코드 내에 학습된 모델 주소를 입력하고 실행

4.4 파라미터 옵션

- train.py
version: 모델을 저장할 때 지칭할 이름 설정

epochs: 모델 학습 시에 반복할 epoch 수
 lr: 모델 학습 시에 사용할 learning rate
 b: 모델 학습 시에 사용할 batch size
 seed: 학습 전 weight를 initialization 할 때 필요한 random seed

5. 베이스라인 설명

5.1 개요

서브태스크 1의 목적은 패션 이미지의 감성 속성들을 추출하고 이에 대한 정확도를 평가함에 있다. 사용하게 될 감성 속성은 일상성, 성, 장식성 총 3가지로 각각의 속성들은 각기 다른 라벨들을 포함하고 있다. 가령, 일상성은 실내복 스타일, 가벼운 외출 스타일, 격식차린 스타일 등 해당 패션이 어떤 상황에 주로 입는지에 따라 분류된 속성이다. 해당 태스크에서는 3개의 감성특징들에 대해 각각 속성 분류를 진행하고, 그에 따른 성능을 평가한다. 이때 성능 평가 metric은 top-1 accuracy(%)가 된다. 전체 성능은 3개의 개별적인 성능을 평균낸 값으로 한다.

통상적으로 딥러닝을 기반으로 한 속성 분류 모델은 인공신경망, 특히 Convolutional layer들로 구성된 CNN을 통해 추출된 feature를 사용한다. Backbone network로 사용되는 CNN은 이미지의 전반적인 특징들을 잘 추출한다고 알려진 ResNet을 사용한다. 얻어진 feature는 fully-connected layer들로 이루어진 classifier를 통과하여 각각의 속성들을 얻게 된다. 해당 구조에 대한 예시는 (그림 5)와 같다.

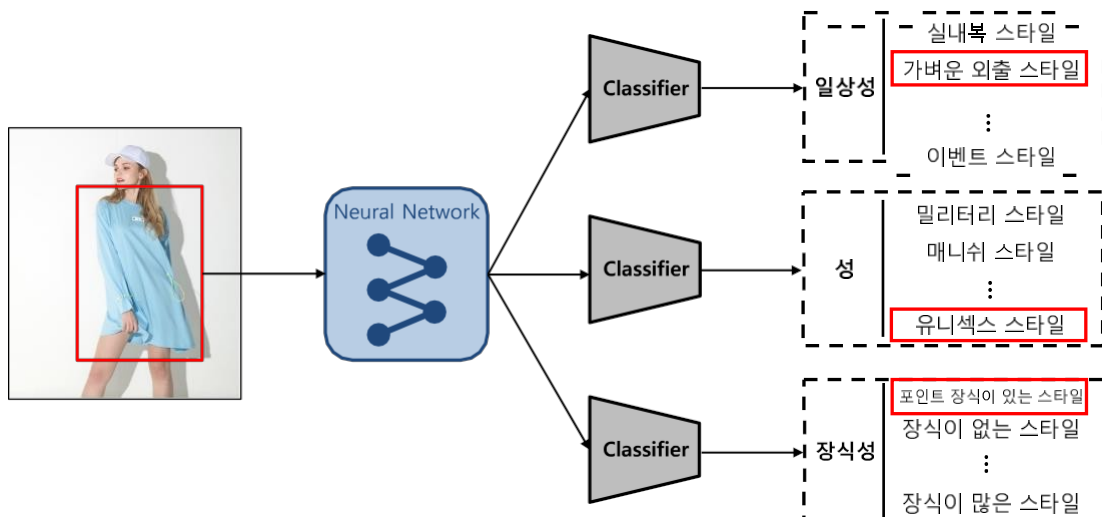


그림 5. 딥러닝 기반의 패션 속성 추출 모델 예시

5.2 학습 및 평가

- <train.py>에 있는 “main” 함수를 사용하여 모델을 학습한다. 미리 설정한 “version” 에 해당하는 이름으로 폴더가 생성되며, 폴더 내에 일정 주기별로 학습된 weight 가 저장된다. “epochs”, “lr”, 그리고 “batch size” 파라미터들을 변경해가며 학습의 정도를 조절할 수 있다.
- <test.py>에 있는 “main” 함수를 사용하여 모델을 평가할 수 있다. 학습된 weight 의 주소를 “trained_weights” 변수 정의 시에 사용되는 함수인 “torch.load(~)”의 입력 인자로 전달한다. 실행 시에 top-1 accuracy 및 acca 에 대한 값이 출력된다.
- <test.py>를 통해 얻어진 일상성, 성, 장식성에 대한 각각의 top-1 accuracy 를 평균내어서 전체 top-1 accuracy 를 구한다.
- 제출시에는 <predict.py>를 통해 제출 결과를 생성하게끔 한다.
이하 그림의 csv 양식 그대로에서 추론 대상 항목(붉은색)에 추론 결과 값을 채워서 제출. (컬럼명 Daily, Gender, Embellishment 유지 필수!)

image_name	BBox_xmin	BBox_ymin	BBox_xmax	BBox_ymax	Daily	Gender	Embellishment	Split
1209/BO00006-1.jpg	2	38	678	870	-1	-1	-1	test
1209/BO00015-1.jpg	93	193	446	614	-1	-1	-1	test
1209/BO00019-1.jpg	59	67	451	648	-1	-1	-1	test
1209/BO00027-1.jpg	127	108	453	561	-1	-1	-1	test
1209/BO00028-1.jpg	59	180	629	480	-1	-1	-1	test
1209/BO00029-1.jpg	23	120	413	530	-1	-1	-1	test
1209/BO00034-1.jpg	0	91	712	727	-1	-1	-1	test
1209/BO00037-1.jpg	207	1	647	717	-1	-1	-1	test
1209/BO00038-1.jpg	156	5	620	412	-1	-1	-1	test
1209/BO00047-1.jpg	36	4	758	707	-1	-1	-1	test
1209/BO00049-1.jpg	157	24	565	548	-1	-1	-1	test
1209/BO00062-1.jpg	56	18	584	650	-1	-1	-1	test
1209/BO00066-1.jpg	1	50	758	702	-1	-1	-1	test
1209/BO00073-1.jpg	178	7	638	683	-1	-1	-1	test
1209/BO00079-1.jpg	16	147	384	531	-1	-1	-1	test

6. 라이선스

- 본 소프트웨어는 MIT 라이선스(<https://opensource.org/licenses/MIT>)를 따라야 함
- 타 오픈소스 SW 활용 시 해당 오픈소스 SW 라이선스에서 요구하는 라이선스 준수 의무를 이행해야 함

7. 연락처

담당자: 박민호

E-mail: roger618@etri.re.kr