

소 속 시각지능연구실	문서명 02_README.pdf	작성일 2023.04.04.	전체 페이지수 7
	작성자 박민호	제 목 챌린지 서브태스크 2 베이스라인 설명	문서관리자 송화전

챌린지 서브태스크 2

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.....	5
4. 실행	5
4.1 실행 환경	5
4.2 설치된 라이브러리	6
4.3 실행 방법	6
4.4 파라미터 옵션.....	6
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 “2023 ETRI 자율성장 인공지능 경진대회”에서 서브태스크 2 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 그림 1)과 같음.

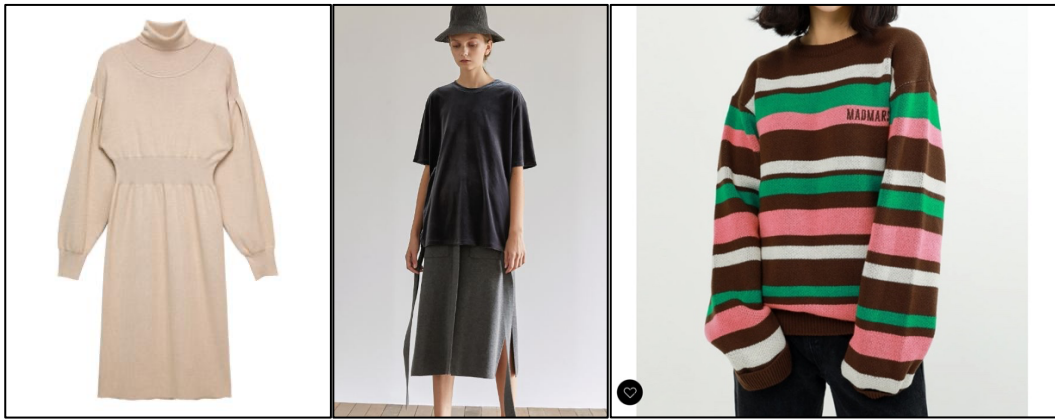


그림 1. 패션 이미지 예시

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

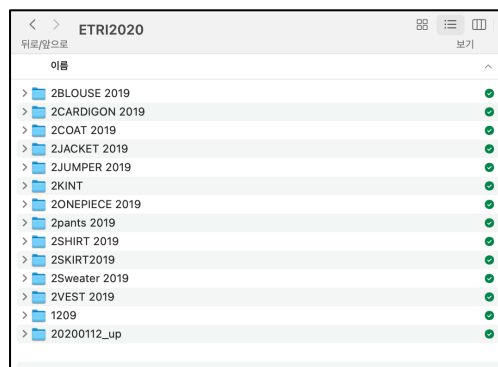


그림 2. 이미지 DB 폴더 구성

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)

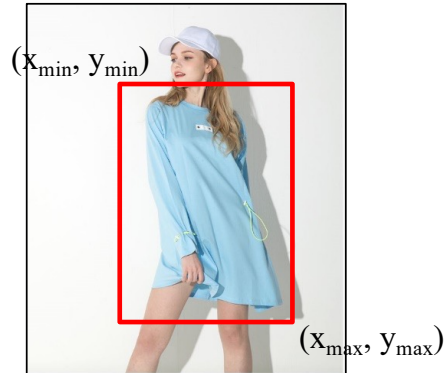


그림 3. 바운딩 박스를 포함한 이미지 예시

- 라벨 정보 DB 예시는 (그림 4)와 같음.

	A	B	C	D	E	F
1	image_name	BBox_xmin	BBox_ymin	BBox_xmax	BBox_ymax	Color
2	1209/BO00001-1.jpg	45	197	313	393	18
3	1209/BO00002-1.jpg	117	114	476	556	16
4	1209/BO00003-1.jpg	62	212	372	544	16
5	1209/BO00004-1.jpg	159	17	484	395	16
6	1209/BO00005-1.jpg	51	54	561	441	16
7	1209/BO00006-1.jpg	2	38	678	870	18
8	1209/BO00007-1.jpg	167	184	435	444	5
9	1209/BO00008-1.jpg	116	205	404	425	16
10	1209/BO00009-1.jpg	124	239	380	423	18
11	1209/BO00010-1.jpg	22	104	499	418	5
12	1209/BO00011-1.jpg	189	231	357	463	18
13	1209/BO00012-1.jpg	106	240	426	641	6
14	1209/BO00013-1.jpg	118	171	402	583	18
15	1209/BO00014-1.jpg	30	59	611	777	18
16	1209/BO00015-1.jpg	93	193	446	614	18
17	1209/BO00016-1.jpg	46	75	574	654	6
18	1209/BO00017-1.jpg	125	144	404	405	15
19	1209/BO00019-1.jpg	59	67	451	648	16
20	1209/BO00020-1.jpg	14	6	702	725	17
21	1209/BO00021-1.jpg	107	40	395	435	16
22	1209/BO00022-1.jpg	79	192	427	656	18
23	1209/BO00023-1.jpg	84	74	361	386	18
24	1209/BO00024-1.jpg	57	57	385	617	17
25	1209/BO00025-1.jpg	103	154	331	474	10
26	1209/BO00026-1.jpg	55	82	336	503	18
27	1209/BO00027-1.jpg	127	108	453	561	6
28	1209/BO00028-1.jpg	59	180	629	480	7
29	1209/BO00029-1.jpg	23	120	413	530	18
30	1209/BO00030-1.jpg	37	13	518	828	16
31	1209/BO00033-1.jpg	146	0	498	499	18

그림 4. 라벨 정보 DB 예시

- 라벨 별 데이터 수는 불균등하게 구성되어 있으며, 데이터 수가 가장 많은 라벨과 가장 적은 라벨의 이미지 수 차이는 약 100배 차이가 남. 라벨 별 데이터 수 분포는 (그림 5)와 같음.

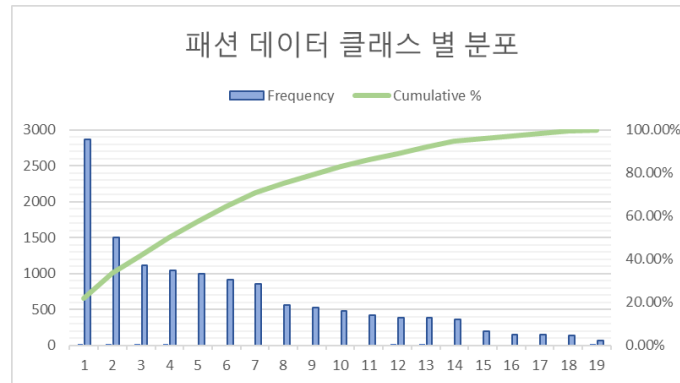


그림 5. 패션 데이터 라벨 별 데이터 수 분포

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
 Python 3.6
 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 2의 목적은 패션 이미지 분류에서 가장 큰 문제점 중 하나인 불균형 데이터 분류 문제를 해결하기 위함이다. 패션 데이터의 특징 카테고리 중 불균형 정도가 심한 카테고리 중 하나인 컬러 특징을 사용한다. 성능 평가 metric은 top-1 accuracy(%)이다.

베이스라인 모델에서는 딥러닝 기반의 classification 모델에서 주로 사용하는 구조인 ResNet을 사용하였다. ResNet에서 얻어진 feature를 fully-connected layer들로 구성된 classifier에 통과시키며 원하는 색채 특징 라벨을 얻는다.

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 aca에 대한 값이 출력된다. 여기서 얻어진 top-1 accuracy로 성능 비교를 하게 된다.

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