롯데정보통신 Vision AI 경진대회

다량의 상품에 대해 각 상품을 특정하는 이미지 분류

롯데정보통신

Contents

01 Storyline

Simple classification

Sample train dataset

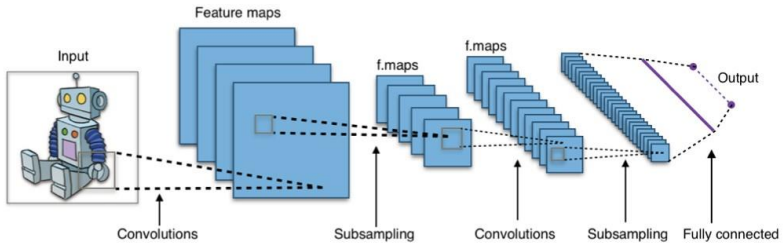
어성초!!

간단한 이미지 분류 문제

01 Storyline

Simple classification

CNN



Convolution Neural Network

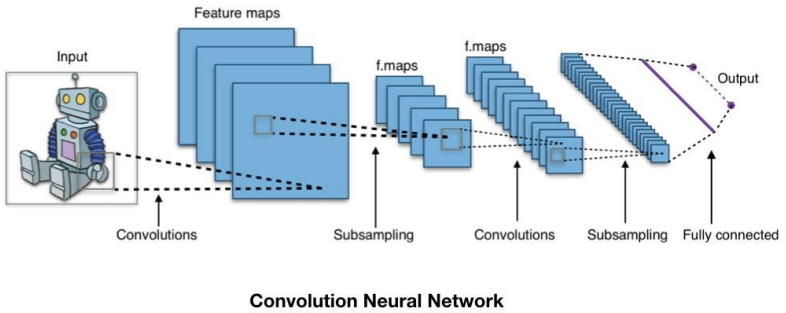
Simple CNN classification??

그저 그런 Validation score

01 Storyline

With label smoothing

CNN



Convolution Neural Network

Simple CNN classification??

Label smoothing

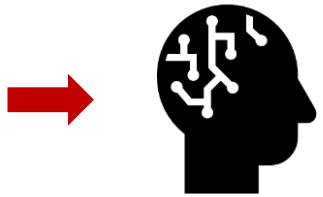
Validation score → 0.99

01 Storyline

With label smoothing

Why label smoothing??

Input



CNN model

환타!

일반적인 이미지 분류는
이미지가 들어오면 해당 클래스를 판단

01 Storyline

With label smoothing

Why label smoothing??

Input

CNN model

환타이긴 한데
핫식스일 수 도??

Label smoothing을 도입하게 된다면

01 Storyline

With label smoothing

Why label smoothing??

≠

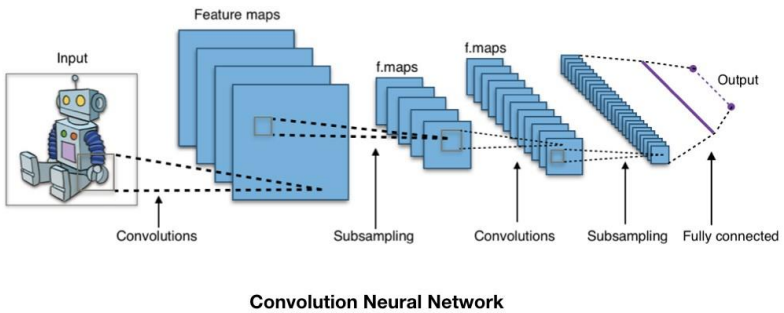
서로 다른 class 이지만 서로 모양은 유사

Label smoothing으로 인해 다른 class임에도
제품의 모양기반으로 feature를
만들어 낼 수 있을 것이라고 기대

01 Storyline

With label smoothing

CNN



Convolution Neural Network

Simple CNN classification??

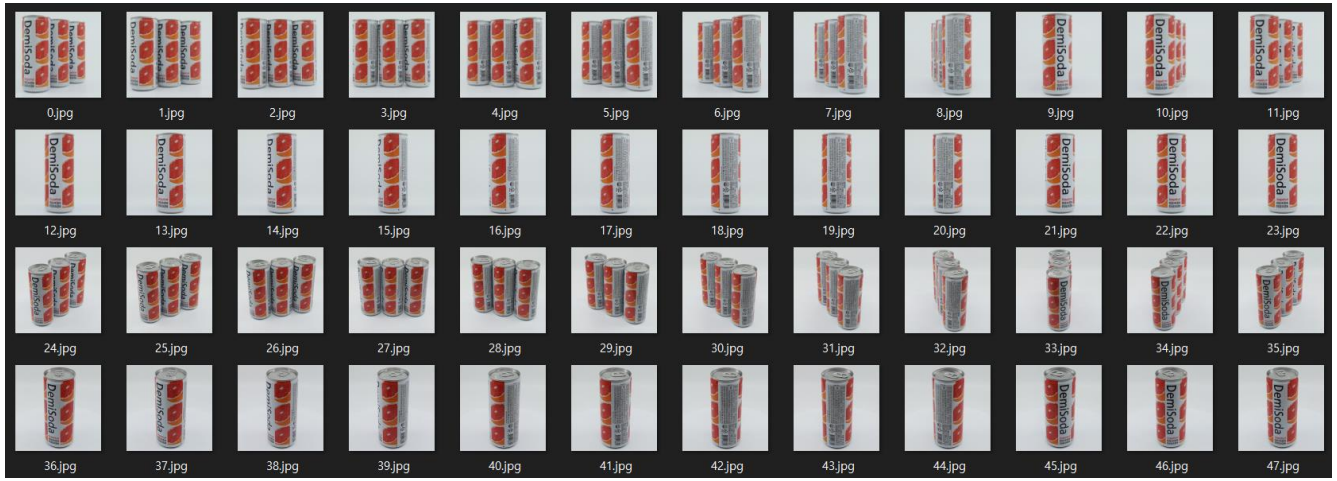
+ Label smoothing = **BUT** Leader board → 0.75

Validation score가 높게 나온 반면
Leader board score는 생각보다 낮음

01 Storyline

With label smoothing

Why??



1. Class별 데이터가 적음(class별 45건)

2. Train에서는 보기 힘든 데이터가 Test dataset에 포함

01 Storyline

With label smoothing

Triplet loss

Few shot learning의 방법 중 하나인
triplet loss를 사용하게 된다면
데이터들의 pair들로 feature들을 생성가능

01 Storyline

With label smoothing

Arcface loss

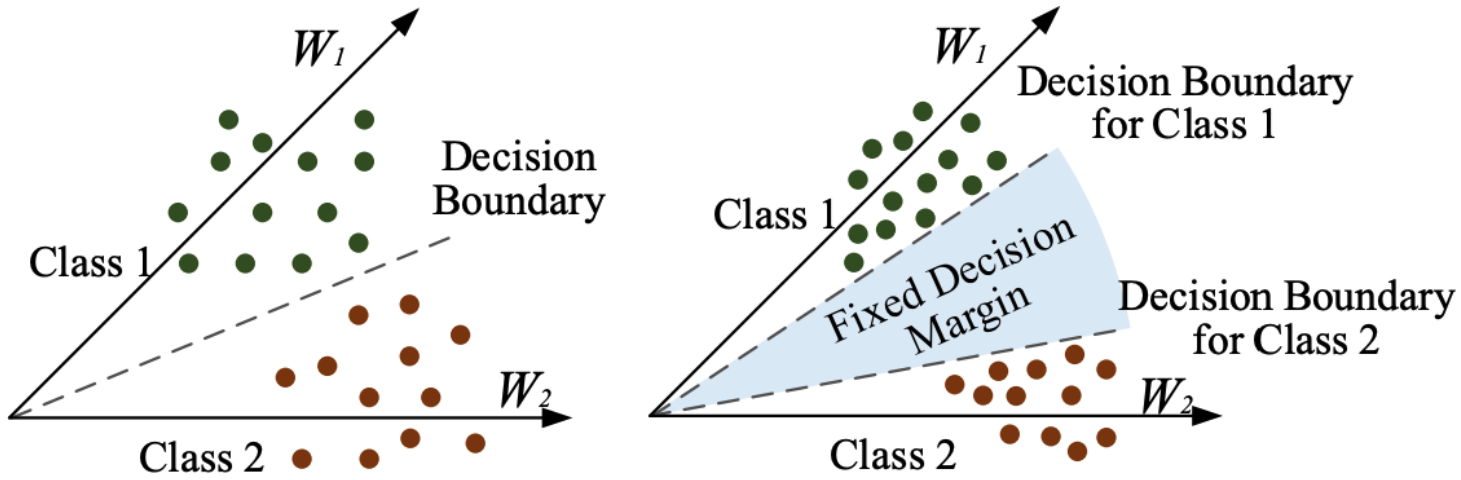
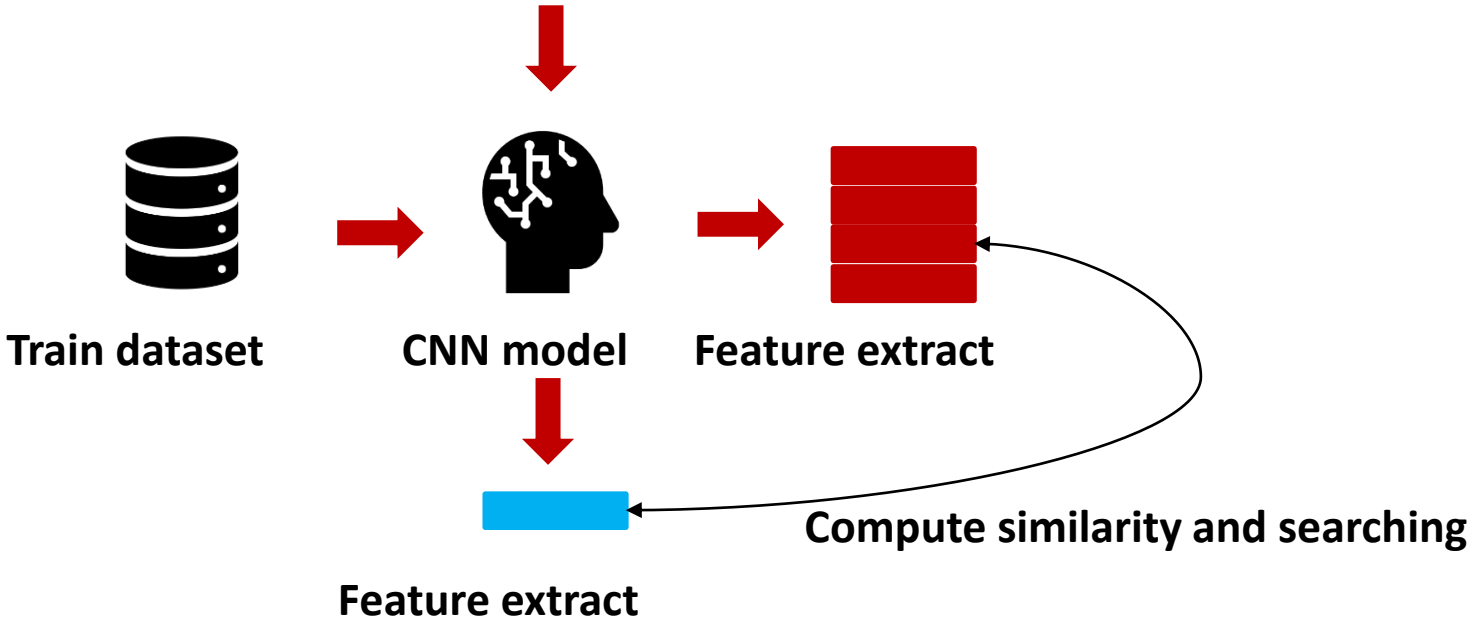


image retrieval 문제에서 triplet loss 보다
arcface loss가 더 좋은 성능을 보임

출처 : ArcFace: Additive Angular Margin Loss for Deep Face Recognition

02 Methods

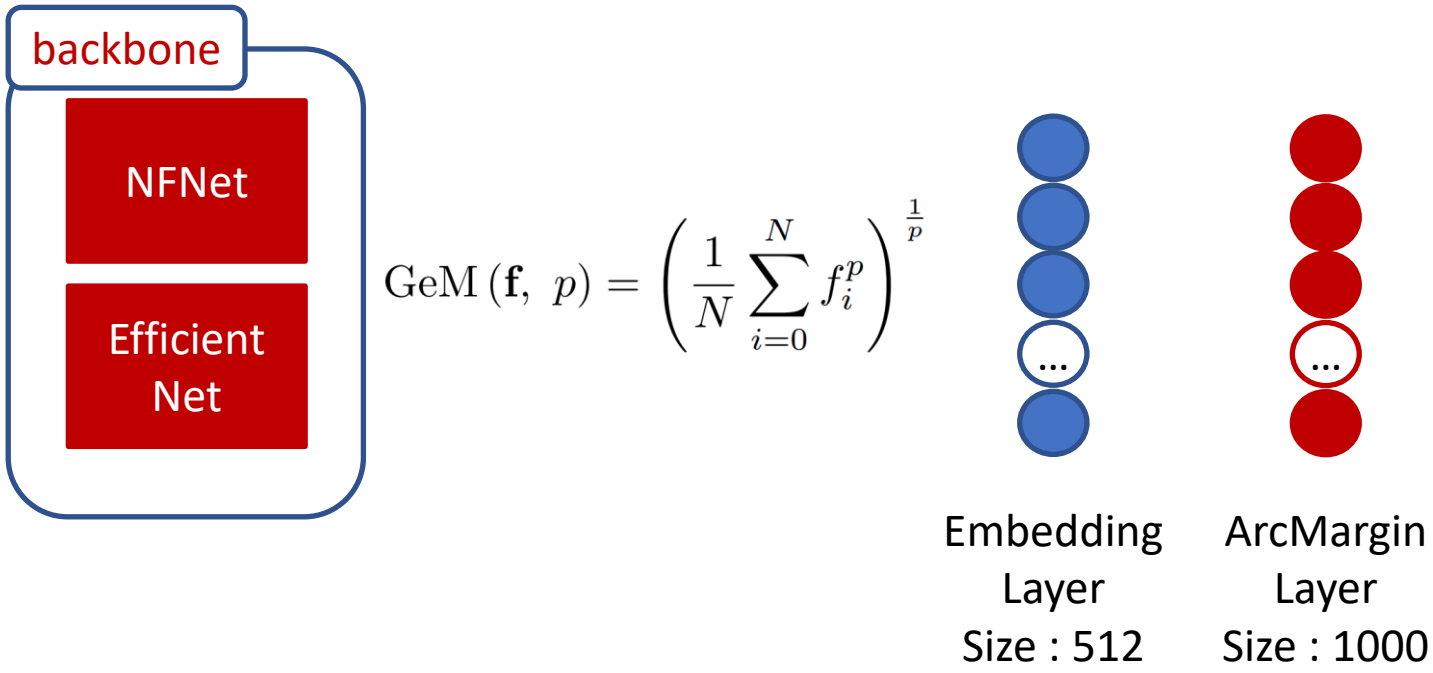
Image retrieval



1. Train Dataset들을 CNN을 통해 Feature extract하여 벡터 확보
2. Test data가 들어오면 CNN을 이용하여 벡터 확보
3. Test 벡터를 train 벡터들과 유사도 기반으로 class 예측

02 Methods

Model architecture



Backbone network로 NFNet과 efficientNet을 이용
Pooling layer로 GeM Pooling 이용
Extract하는 vector size는 512
Output layer로 Arcface loss를 위한 ArcMargin layer 사용

02 Methods

Loss

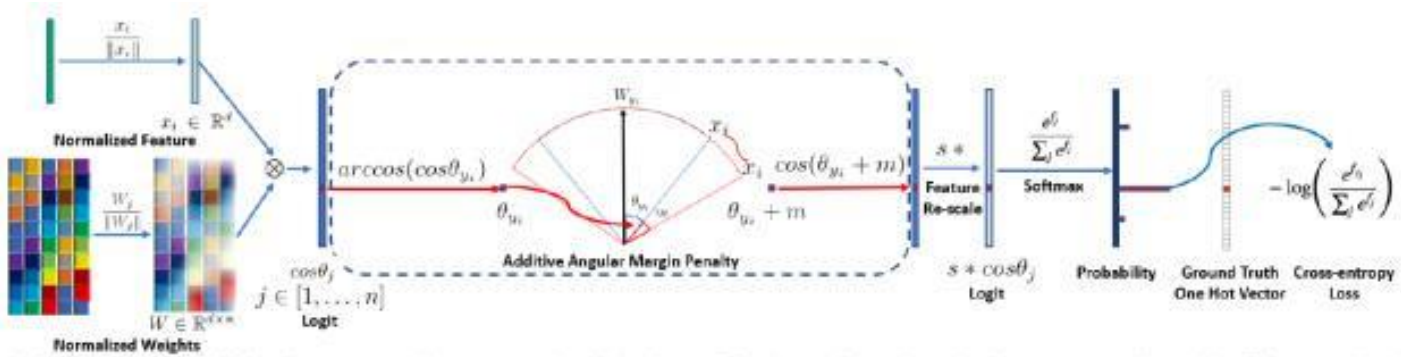


Figure 2. Training a DCNN for face recognition supervised by the ArcFace loss. Based on the feature x_i and weight W normalisation, we get the $\cos\theta_j$ (logit) for each class as $W_j^T x_i$. We calculate the $arccos\theta_{y_i}$ and get the angle between the feature x_i and the ground truth weight W_{y_i}. In fact, W_j provides a kind of centre for each class. Then, we add an angular margin penalty m on the target (ground truth) angle θ_{y_i}. After that, we calculate $\cos(\theta_{y_i} + m)$ and multiply all logits by the feature scale s. The logits then go through the softmax function and contribute to the cross entropy loss.

$$L = -\frac{1}{N}\sum_{i=1}^{N}\log\left(\frac{e^{s(\cos(\theta_{y_i}+m))}}{e^{s(\cos(\theta_{y_i}+m))} + \sum_{j=1,j\neq y+i}^{n} e^{s\cos(\theta_j)}}\right)$$

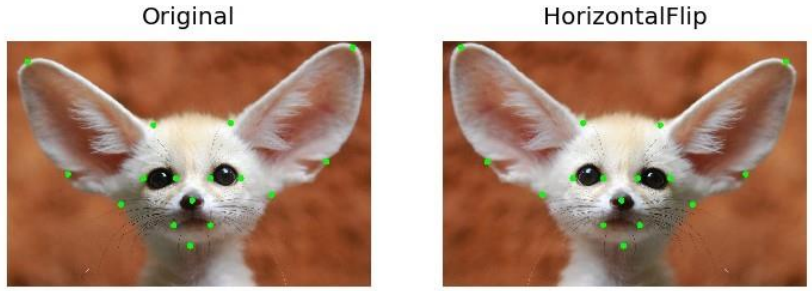
출처 : ArcFace: Additive Angular Margin Loss for Deep Face Recognition

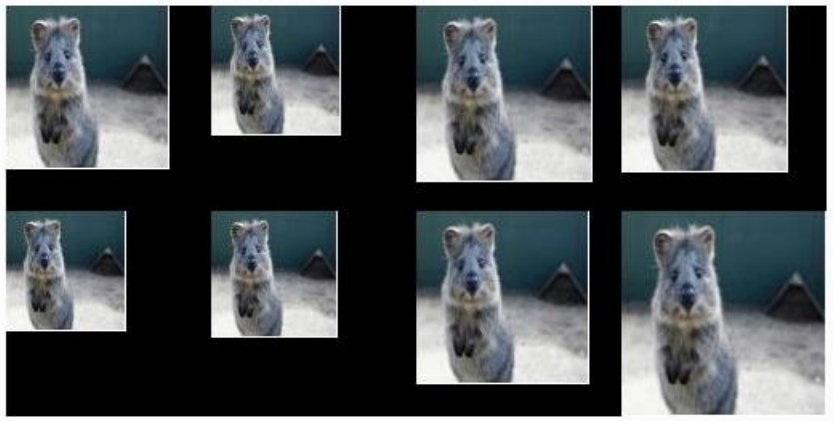
02 Methods

Train augmentation

Image Compression

Cutout





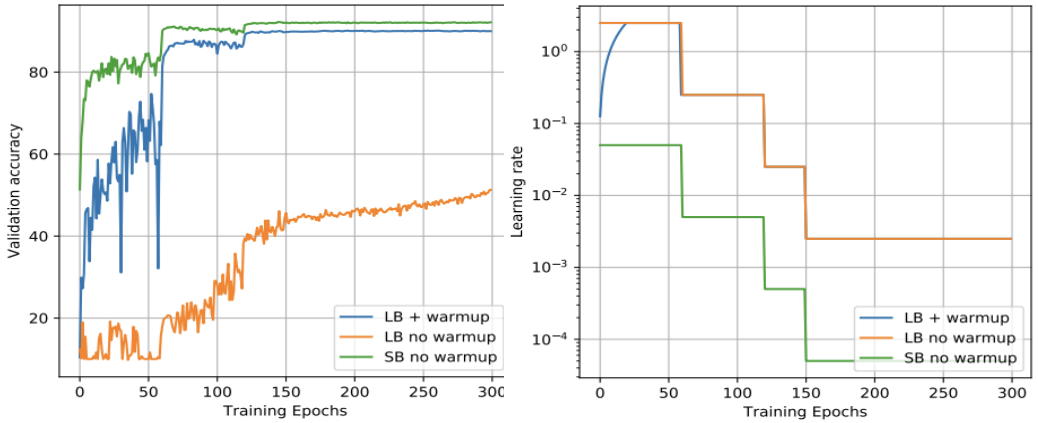
Shift Scale Rotate

Train augmentation으로
image compression
Cutout
Horizontal Flip
Shift scale rotate
적용

02 Methods

Learning rate scheduler

Warm-up 스케줄링 + cosine anneling

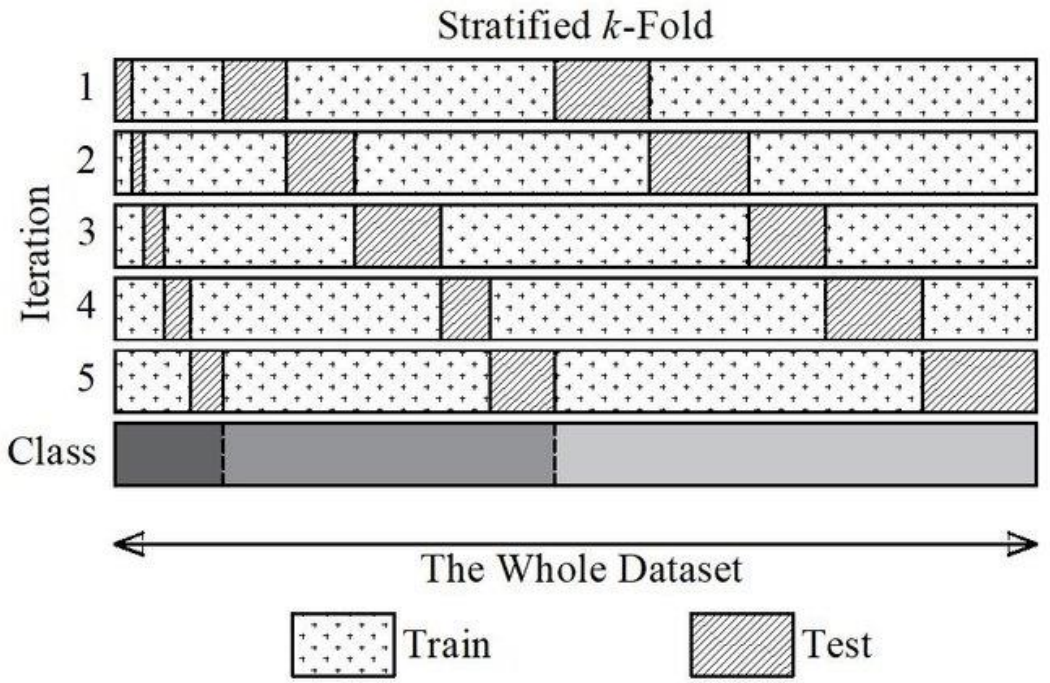


다양한 스케줄링 방법 중 warm-up scheduling + cosine annealing 기법을 사용

출처 : A closer look at Deep learning heuristics: Learning rate starts, warmup and distillation (2018), Richard Socher et al

02 Methods

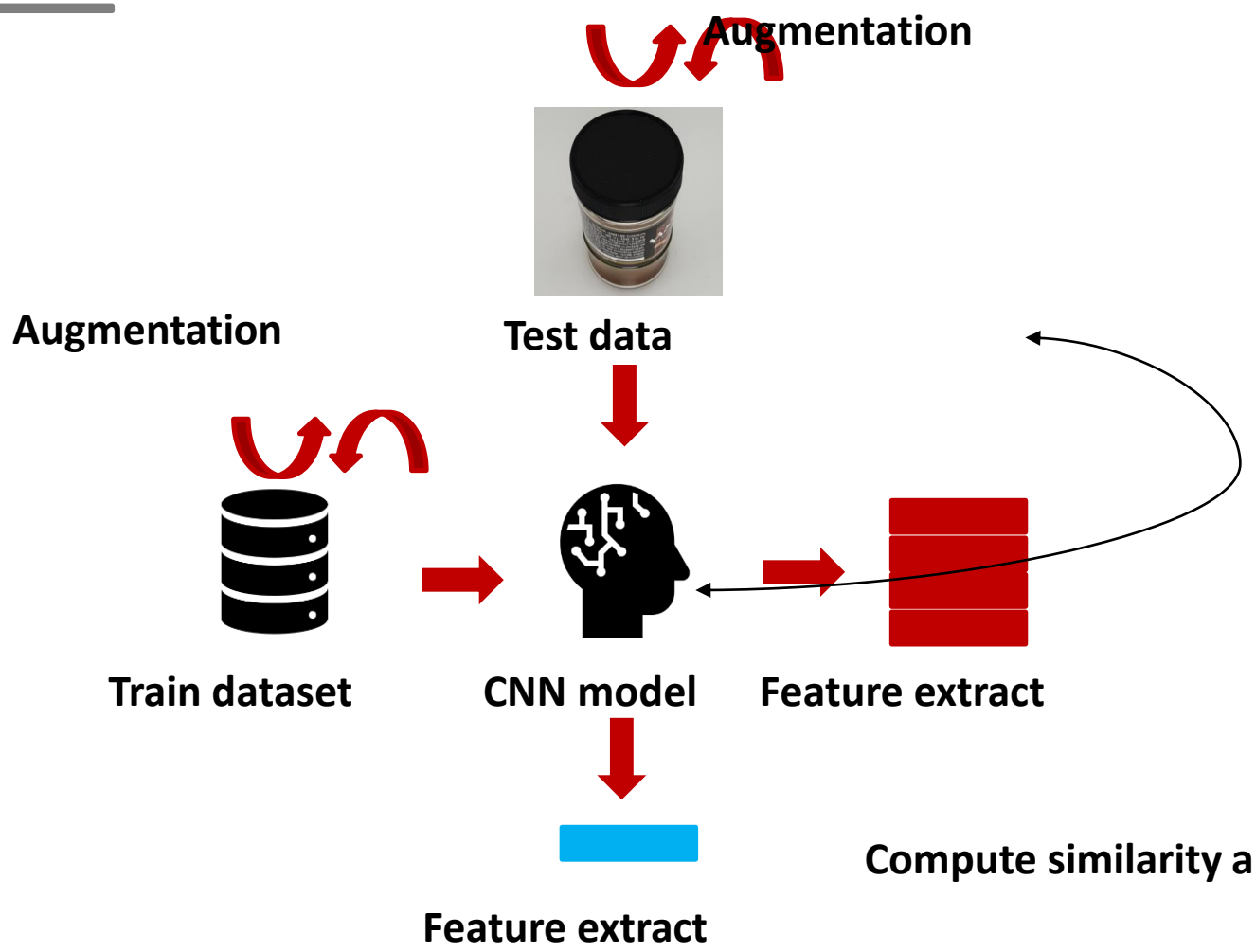
Validation strategy



Validation set으로 class를 stratified하게 추출
NFNet, EfficientNet 각 5개의 fold 생성

02 Methods

TTA



Test time에서도 Augmentation을 적용하여 모델에게 다양한 데이터를 보여줘 Inference를 진행

03 Results

개요 데이터 대회규칙 리더보드 포럼 종료

실시간 리더보드

- 실시간 리더보드 순위는 Validation 측정 결과만을 반영한 것으로, 최종 순위와 다를 수 있습니다.
- 최종 순위는 별도의 검증절차를 통해 결정됨을 참고하여 주세요.

순위	이름	스코어	제출날짜
1	skyblue93**	95.436	2021 . 03 . 25 (23 : 25 : 06)
2	hakddal7**	95.150	2021 . 03 . 26 (17 : 52 : 22)
3	jeong59**	94.985	2021 . 03 . 26 (17 : 44 : 49)

감사합니다.