

롯데정보통신 Vision AI 경진대회 솔루션 공유

김 명 환

chadool116@naver.com

목차

1. 데이터 소개 및 전처리
2. 모델 구축 및 검증
3. 결과 및 결론

1. 데이터 소개 및 전처리 – 데이터 소개

- 학습용 데이터 48,000장 (1,000 Class * 48개)



- 시험용 데이터 72,000장 (1,000 Class * 72개)

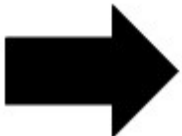


1. 0도, 30도(pitch)로 구성 되어 있는 Train 데이터를 활용하여 AI 모델 학습
2. 0도, 30도, 60도(pitch)로 구성된 Test 데이터 상품 이미지를 분류
3. 학습용 데이터에서 찾을 수 없는 특징이 Test 데이터셋에 존재

Test 데이터가 Train 데이터보다 많으므로 데이터를 증강시켜 학습할 필요가 있다고 판단하였음

1. 데이터 소개 및 전처리 – 전처리 작업

1. 학습 효과를 높이기 위한 데이터 정규화 진행
 - 0 ~ 255 분포값을 가지는 이미지 정보를 0 ~ 1 사이의 값으로 정규화

<pre>array([[208, 215, 218], [208, 215, 218], [208, 215, 218], ..., [205, 212, 215], [205, 212, 215], [205, 212, 215]], [[208, 215, 218], [208, 215, 218], [208, 215, 218],</pre>		<pre>array([[0.85490196, 0.84313725, 0.81568627], [0.85490196, 0.84313725, 0.81568627], [0.85490196, 0.84313725, 0.81568627], ..., [0.84313725, 0.83137255, 0.80392157], [0.84313725, 0.83137255, 0.80392157], [0.84313725, 0.83137255, 0.80392157]], [[0.85490196, 0.84313725, 0.81568627], [0.85490196, 0.84313725, 0.81568627], [0.85490196, 0.84313725, 0.81568627],</pre>
---	---	--

1. 데이터 소개 및 전처리 – 전처리 작업

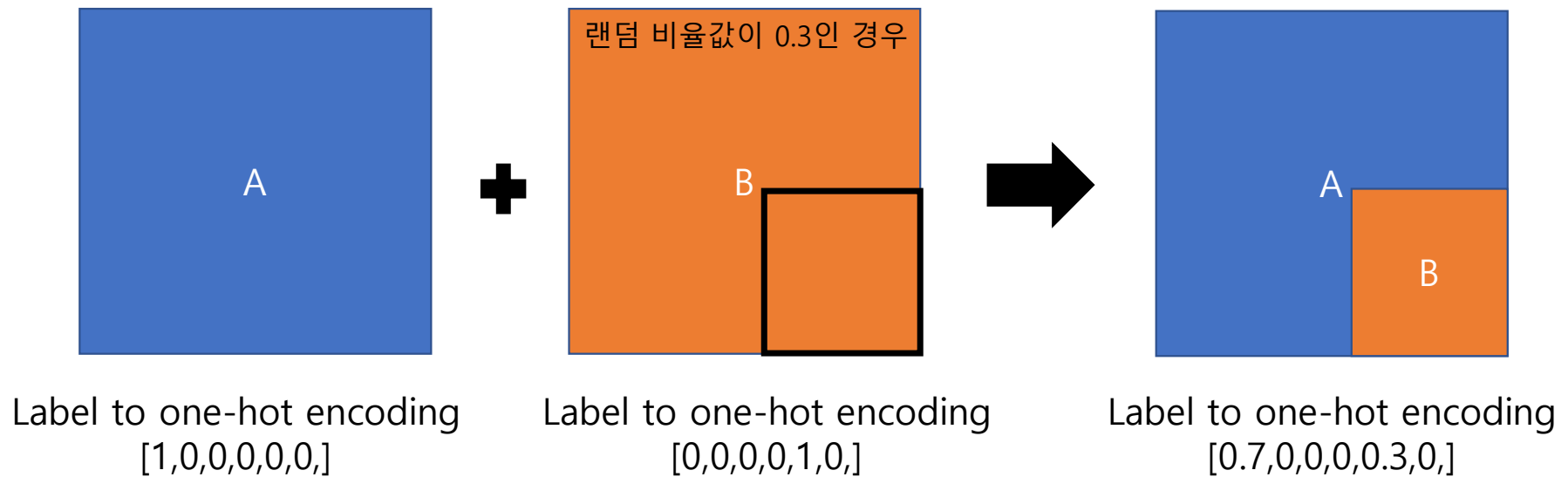
2. Augmentation(수평, 밝기, 명암, 확대조정) 적용



1. 데이터 소개 및 전처리 – 전처리 작업

3. Cutmix 적용

- 랜덤비율만큼 바운딩박스 영역을 추출하여 2개의 이미지를 하나의 이미지로 합성



1. 데이터 소개 및 전처리 – 전처리 작업

3. Cutmix 적용

- 랜덤비율만큼 바운딩박스 영역을 추출하여 2개의 이미지를 하나의 이미지로 합성



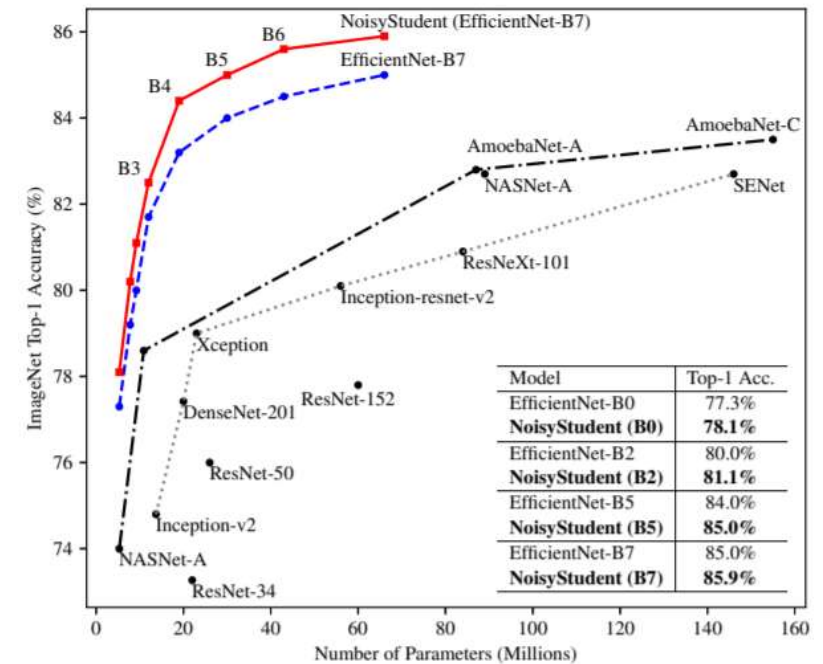
2. 모델 구축 및 검증

1. 사용 모델 및 파라미터 정보

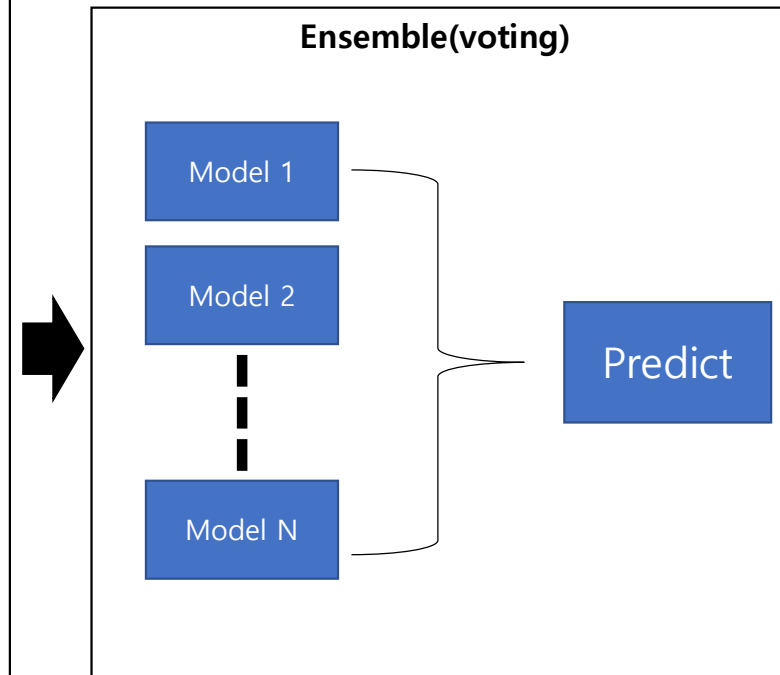
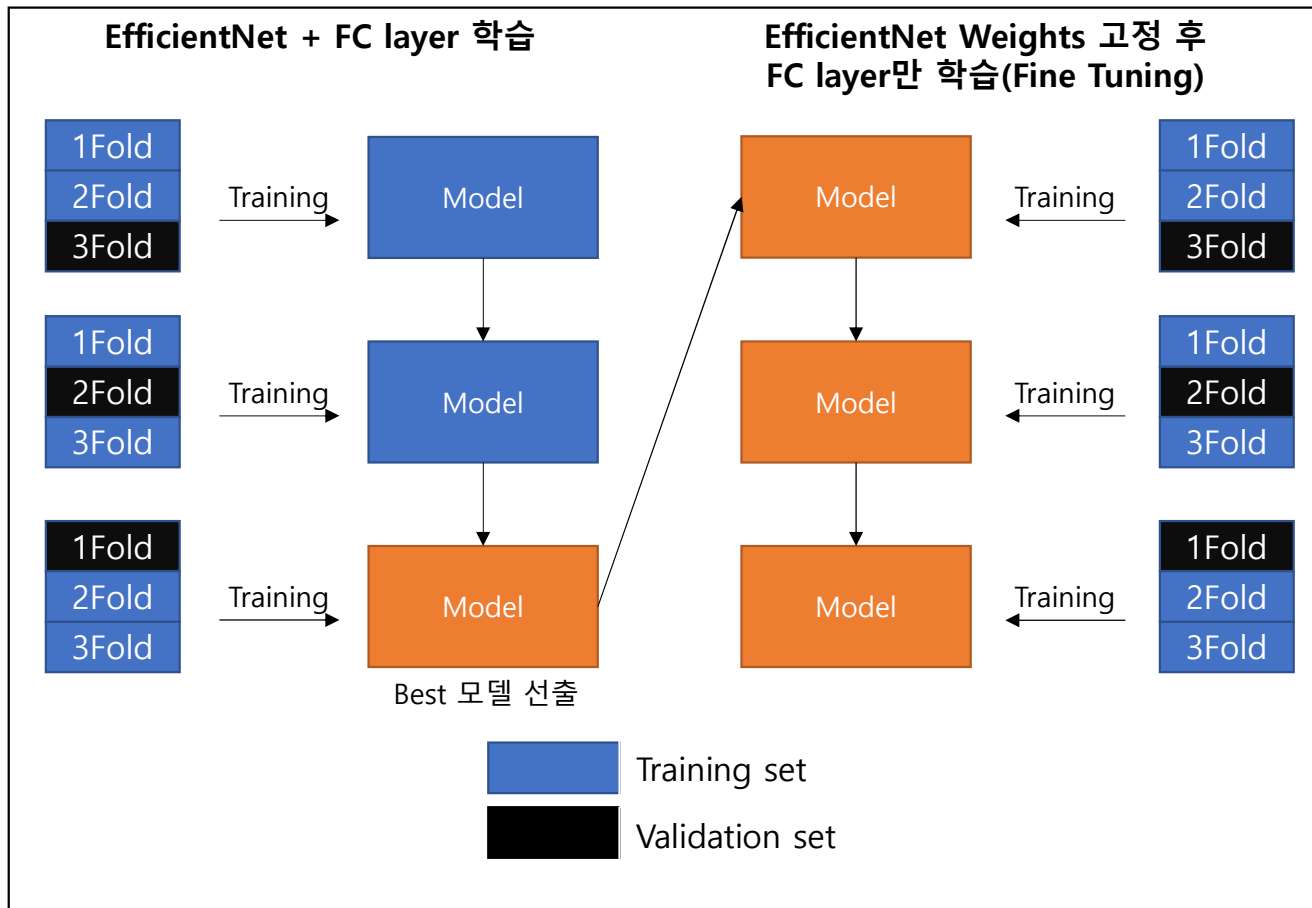
- Base모델: EfficientNet(B0~b6) + Fully Connected Layer(FC)
- 사용한 Pre-trained 모델: noisy-student
- 학습 이미지 사이즈: 200 ~ 512
- Optimizer : Adam
- Loss function : Categorical Cross-Entropy
- Learning Rate Schedule : Warmup

2. 학습 정보

- 48,000개의 학습 데이터를 3개의 Fold로 나누어 학습
 - Train data: 32,000
 - Validation data: 16,000
- Fold별 2회(총6회) 모델 학습 후 Fine Tuning 기법을 이용하여 CNN Layer Weights 고정, Fully Connected Layer 만 추가 학습 진행
- 학습된 모델들을 종합하여 Ensemble(Voting)기법을 활용하여 클래스 분류



2. 모델 구축 및 검증



3. 결과 및 결론

- 테스트 환경
 - Google Colab
 - 사용 GPU: Tesla K80(12G RAM)
- EfficientNet 모델이 깊어질수록(B0 -> B6) 학습 대비 정확도가 떨어졌으며 B1, B2에서 높은 정확도를 달성
- 이미지 resize(200 ~ 512) 조정에 따른 학습결과 원본 이미지 사이즈(256)에서 높은 정확도를 달성
- Augmentation + Cutmix 기법을 적용하여 가장 높은 정확도(EfnB2, LB : 91.092)를 달성할 수 있었음
- 학습된 모델들을 Ensemble하여 예측한 결과 리더보드 기준 93.421점을 달성할 수 있었음

Augmentation	Cutmix	LB Score
X	X	50 ~ 80
O	X	78 ~ 88
O	O	85 ~ 91

전처리 여부에 따른 LB Score 분포

EfficientNet	Max LB Score
B0	85.569
B1	88.699
B2	87.815
B3	87.587
B4	84.196
B5	78
B6	79

Augmentation 사용여부에 따른
Max LB Score

EfficientNet	Max LB Score
B1	90.661
B2	91.092
B3	90.835

Augmentation + Cutmix
사용여부에 따른 Max LB Score

3. 결과 및 결론

- 테스트 환경
 - Google Colab
 - 사용 GPU: Tesla K80(12G RAM)
- EfficientNet 모델이 깊어질수록(B0 -> B6) 학습 대비 정확도가 떨어졌으며 B1, B2에서 높은 정확도를 달성
- 이미지 resize(200 ~ 512) 조정에 따른 학습결과 원본 이미지 사이즈(256)에서 높은 정확도를 달성
- Augmentation + Cutmix 기법을 적용하여 가장 높은 정확도(EfnB2, LB : 91.092)를 달성할 수 있었음
- 학습된 모델들을 Ensemble하여 예측한 결과 리더보드 기준 93.421점을 달성할 수 있었음

EfficientNet Model List	Augmentation	CutMix	LB Score	Ensemble LB Score
B0	O	X	85.569	93.421
B1	O	X	88.699	
B2	O	X	87.815	
B3	O	X	87.587	
B1	O	O	90.661	
B2	O	O	91.092	
B3	O	O	90.835	

Ensemble 적용에 따른 LB Score

순위	이름	스코어
1	skyblue93**	95.436
2	hakddal7**	95.150
3	jeong59**	94.985
4	thisisir**	94.761
5	chadool1**	93.421
6	maiho**	92.796
7	boysbo**	92.575
8	edshkim**	92.083

감사합니다!