

# **딥러닝을 통한 두피 진단**과 **두피케어 제품 추천 서비스 개발**

**멀티캠퍼스 AI SW 개발자 부트캠프** FINAL PROJECT

2조 DeepRoot



김 성 환ㆍ김 준 호ㆍ이 혜 진ㆍ전 민 정

# **CONTENTS**

01

프로젝트 소개

목적

02

서비스 소개

소개

제품추천 방식

03

데이터 소개

Al Hub

올리브영

04

모델링

제한사항

개선 과정

결과

05

한계

# **CONTENTS**

01

프로젝트 소개

목적

02

서비스 소개

UTT -E--EL HEA

03

데이터 소개

Al Hub

올리브영

04

모델링

제한사형

개선 과정

결괴

0.5

하계

# 1. 프로젝트 소개

1) 목적: 두피 관리에 대한 소비자 관심 증가

# 약 200억 달러

2024 - 2031년 예상 시장 규모

약 126억 달러

23' 두피 관리 기기 글로 벌 시장 규모 >

약 100억 달러

23' 장난감 글로벌 시장 규모

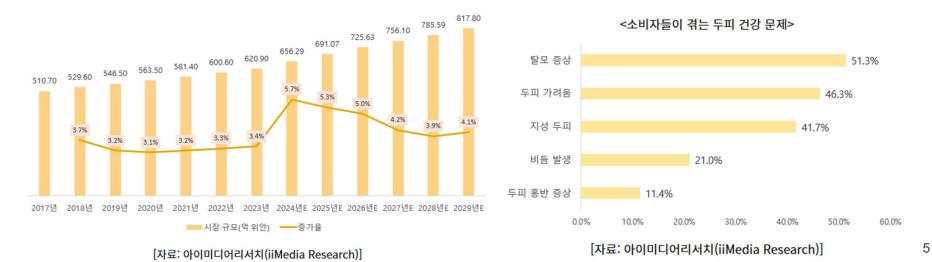
출처: Verified Market Research, Bonafide Research, 2023

출처: Verified Market Research, Bonafide Research, 2023

# 1. 프로젝트 소개

### 1) 목적: 두피 관리에 대한 소비자 관심 증가





# **CONTENTS**

01

**프로젝트 소개** <sup>목적</sup> 02

서비스 소개

소개

제품추천 방식

03

데이터 소개

Al Hub

올리브영

04

모델링

제한사항

개선 과정

결괴

05

하겨



두피 관리는 병원에 가서 받으려면 대기 시간이 길고, 진단을 위해 여러 장비로 사진을 찍느라 움직여 야 하며, 의사의 설명을 듣는 것도 번거로울 수 있습니다. 이러한 과정이 귀찮고 어렵게 느껴지기 때문에, 많은 분들이 두피 관리에 소홀해지기 쉽습니다.

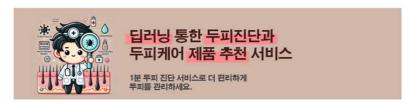
그래서 저희 조는 CNN 아키텍쳐인 Efficientnet으로 학습시킨 모델을 활용하여 두피 이미지를 분석하고 증상을 파악합니다. 이를 바탕으로 이용자의 두피 증상에 최적화된 제품을 추천하며, 이미지 업로드 만으로도 빠르고 쉽게 두피 상태를 점검하고 관리 솔루션을 제공받을 수 있습니다.

이를 통해, 두피 문제를 사전에 감지하여 증상이 악화되기 전에 예방할 수 있으며, 각 사용자의 두피상 태에 맞는 최적의 제품으로 건강한 두피를 유지할 수 있습니다. 또한, 언제 어디서나 두피상태를 점검 할 수 있기에 편리한 경험을 제공합니다.

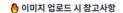
# 2. 서비스 소개

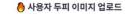
### 1) 소개

### (1) PC 버전

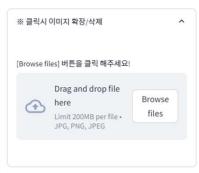


2024.11.13, made by DeepRoot(김성환, 김준호, 이혜진, 전민정)





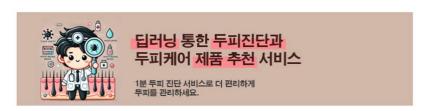




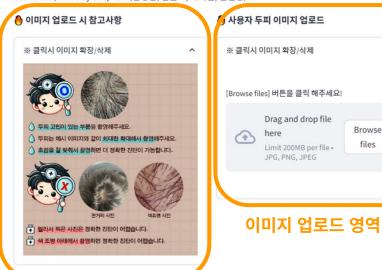
### (2) 모바일 버전



1page



2024.11.13, made by DeepRoot(김성환, 김준호, 이혜진, 전민정)



1page

(1) 이미지 업로드 시 참고사항

- 정확한 진단을 위한 두피 촬영 방법 안내

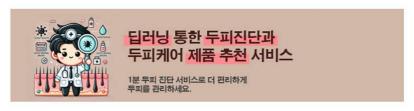
- 색의 영향을 많이 받는 모낭사이홍반 증상의 오진단을 방지하기 위한 구체적인 안내 포함

### (2) 사용자 두피 이미지 업로드

- Drag 또는 Drop 방식으로 이미지 업로드
- 이미지 업로드 완료 시, 문구 출력

이미지가 성공적으로 업로드되었습니다.

촬영 방법 안내 영역



2024.11.13, made by DeepRoot(김성환, 김준호, 이혜진, 전민정)

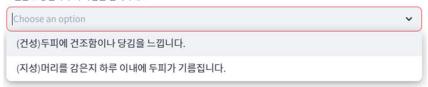
### 설문조사

- 당신의 두피에 대해 알려주세요. 더 정확한 분석에 도움이 됩니다.
- 질문 선택 후 Submit 버튼을 반드시 눌러주세요.



### 설문조사

- 두피 진단의 정확도를 향상시키기 위해 진행.
- 질문 1: 사용자의 두피 타입 파악(지성/건성)
- 두피 타입을 모르는 사용자를 위해, 타입 특징을 설문에 포함
- 질문 1. 당신의 두피 타입을 선택하세요



- 질문 2: 사용자의 고민 또는 진단 질환 파악 ※ 6가지 증상 중 복수 선택 가능

미세각질

비듬

모낭사이홍반

모낭홍반농포

피지과다

탈모

- 질문 3: 추천 페이지에서 사용자가 원하는 제품 카테고리 선택

※ 7가지 제품 카테고리 중 복수 선택 가능

린스/컨디셔너 샴푸바/드라이샴푸

샴푸

헤어오일/헤어세럼

헤어워터

두피팩/스케일러

헤어토닉/두피토닉

10



2024.11.13, made by DeepRoot(김성환, 김준호, 이혜진, 전민정)

### 설문조사

- 당신의 두피에 대해 알려주세요. 더 정확한 분석에 도움이 됩니다.
- 질문 선택 후 Submit 버튼을 반드시 눌러주세요.



### 설문조사 결과 안내

- 설문 조사 선택지를 요약 제공하여 잘못 선택하지 않았는지 더블 체크.
- 다음 페이지의 내용을 제공하여 페이지 간의 내용 연결.
  - 당신의 두피가 💠 건성 타입의 🔝 탈모,모낭사이홍반 증상이 있다라고 선택하셨습니다.
  - 다음 페이지에서 당신이 💠 건성 타입의 🔝 탈모,모낭사이홍반 증상을 가지고 있는 지 🔠 두피이미 지 분석을 통해 확인해 보도록 하겠습니다.
  - Next 버튼을 눌러주세요!

설문조사 영역

결과 안내 영역

# 2. 서비스 소개

### 1) 소개



### 두피 진단

- 설문조사 결과와 제공받은 이미지를 모델로 예측한 결과를 합하여 더 정확한 진단 제공
- 증상별 두피 이미지를 제공하여 현재 사용자의 두피 상태랑 비슷한 증상이 무엇인지 육안으로 파악 가능
- 모델이 예측한 두피증상을 확률과 함께 제공
- 상태분석 노트에서 앞선 설문결과와 모델 예측 진단 결과를 간략히 요약하고 다음 페이지의 내용과 연결





### 딥러닝 통한 두피진단과 두피케어 제품 추천 서비스

1분 두피 진단 서비스로 더 편리하게 두피를 관리하세요.

2024.11.13, made by DeepRoot(김성환, 김준호, 이혜진, 전민정)

### ● 두피 증상별 원인과 특징 그리고 관리방안

• 당신의 두피 상태인 😇 피지과다의 원인과 특징 그리고 관리방안을 확인하도록 하겠습니다

### 지성 두피 (피지과다)

원인: 호르몬 불균형, 유전적 요인, 스트레스.

특징: 번들거리는 모낭, 모발과 두피가 지루하게 윤기를 띔.

### 관리방안:

- 1. 올바른 클렌징(): 부드러운 클렌징 제품을 사용하여 매일 정기적으로 두피를 세정.
- 2. 적절한 헤어케어🍱 : 지성 특성에 맞는 성분이 함유된 샴푸와 적절한 두피 마사지로 피지 분비를 조절.
- 3. 규칙적인 스킨케어 🎉: 두피 팩이나 토너를 이용하여 두피를 진정시키고 피지 조절에 도움을 줌.
- Next 버튼을 눌러주세요!

Home

Prev

Next

4page

### 두피 증상별 원인과 특징 그리고 관리방안

- 두피 타입과 증상을 고려한 관리 방법 안내
- OpenAI를 통한 서비스 제공
- 두피 상태에 관한 원인과 특징, 관리방안을 ChatGPT에게 질문하고 답을 화면에서 볼 수 있도록 구현





### 추천 제품 목록

- 두피상태와 타입을 <u>재확인</u> 예) 두피상태 : 피지과다 중증, 두피 타입 : 지성
- <u>두피 타입과 증상을 고려하여 선별된 제품 중 별점, 리뷰수</u> <u>순</u>으로 가장 높은 점수를 받은 제품 3개씩 추천
- 제품 이미지를 누르면 올리브영으로 제품을 바로 구매할 수 있도록 함으로 써 **활용도**를 높임
- 제품 설명서에 제품 별로 두피의 타입과 매칭한 성분들을 보여줌으로 써 제품의 선택에 <u>신뢰성</u>을 높임.

# 2. 서비스 소개 2) 제품 추천 방식

a. 두피 타입별 효과 성분 / 출처: 대한화장품협회

지성	건성	민감성	비듬	탈모	지루성
피지, 홍반	각질, 비듬	각질, 비듬	각질, 비듬	탈모	피지, 홍반, 농포
카올린	락토바실러스	락틱애씨드	다이소듐라우레스설포석시네이트	비오틴	케토코나졸
화산재	자이언트켈프 발효여과물	살리살릭애씨드	케토코나졸	카페인	시클로피록스
옥수수전분	글리세린	밴조익애씨드	유황	아미노산	세르타코나졸
알킬폴리글루코사이트	소듐PCA	판테놀	살리실산	판테놀	살리실산
소듐C14-16올레핀설포네이트	히알루론산	비사볼올		니아신아마이드	스쿠알란
소듐메칠코코일타우레이트	세라마이드	마데카소사이드			
살리실릭애씨드	스쿠알란				
클레이					
니아신아마이드					

# 2. 서비스 소개 2) 제품 추천 방식

## b. 제품 성분에 따른 두피 증상 매칭 (예: 샴푸) / 출처: 대한화장품협회

page	product_link	img_link	product_name	star	review_count	ingredient	brand_name	type
1	EC%84%B8&t_number=1	00021083403ko.jpg?l=ko	불기획(+100ml 추가 증정)	4.8	2855	기티에이@1,2-헥산다이올	아로마티카	지성,건성,민감성,탈모
1	EC%84%B8&t_number=2	00021135102ko.jpg?l=ko	네트(+50ml+브러쉬 증정)	4.9	2326	이꽃추출물, 칡뿌리추출물	스팀베이스	듬,지성,건성,민감성,탈모
1	EC%84%B8&t_number=3	00018087212ko.jpg?l=ko	500g (단백질 앰플 샴푸)	4.8	3286	니올, 헥실신남알, 리모넨	어노브	건성,탈모
1	EC%84%B8&t_number=4	00020849201ko.jpg?l=ko	란유래단백질 샴푸 500ml	4.8	411	이드엔피, 포타슘솔베이트	하아르	,건성,지루성,민감성,탈모
1	EC%84%B8&t_number=5	00020826609ko.jpg?l=ko	Dml+100ml 추가증정기획	5.0	4260	드케라틴, 리모넨, 리날물	솔랩	지성,건성,민감성,탈모
1	EC%84%B8&t_number=6	00019255411ko.jpg?l=ko	어 샴푸 500ml 더블 기획	4.8	5944	(3:1), 유칼립투스잎추출물	헤드앤숄더	
1	EC%84%B8&t_number=7	00016447031ko.jpg?l=ko	ML 기획 외 단품 3종 택 1	4.8	1224	이티드레시틴, 카놀라오일	롱테이크	듬,지성,건성,민감성,탈모
1	EC%84%B8&t_number=8	00015247514ko.jpg?l=ko	· 두피강화 (333ml+50ml)	4.7	9686	, 피토스핑고신, 바이오틴	라보에이치	건성,비듬,탈모
1	EC%84%B8&t_number=9	00020127710ko.jpg?l=ko	선 탈모완화 샴푸 1000ml	4.9	1164	인, 포타슘솔베이트, 향료	바이오가	지성,건성,비듬,탈모
1	C%84%B8&t_number=10	00018124725ko.jpg?l=ko	프 샴푸 200ml 듀오 기획	4.8	6531	ɔ 유래 클렌징 성분 함유)	라우쉬	지성
1	C%84%B8&t_number=11	00021119105ko.jpg?l=ko	트) 880ML 단독 기획세트	4.9	9958	.로넬올, 리날물, 제라니올	닥터포헤어	돔,지성,건성,민감성,탈모
1	C%84%B8&t_number=12	00016746404ko.jpg?l=ko	푸 500ml+70ml 증정기획	4.8	2144	·이콜, 잔탄검, 석류추출물	루톤	듬,지성,건성,민감성,탈모
1	C%84%B8&t_number=13	00000962107ko.jpg?l=ko	라이샴푸 오리지널 200ml	4.7	13532	랄,제라니올,리모넨,리날쿨	바티스트	
1	C%84%B8&t_number=14	00017501702ko.jpg?l=ko	트(샴푸 리필 330ml 증정)	4.9	650	코페롣,헥실신남알,리날쿨	츠바키	지성
1	C%84%B8&t_number=15	00019790106ko.jpg?l=ko	리필 기획 (+600ML 리필)	4.8	19525	, 피토스핑고신, 바이오틴	라보에이치	건성,비듬,탈모
1	C%84%B8&t_number=16	00016189027ko.jpg?l=ko	두피 스칼프 샴푸 300ml	4.8	1738	, 동백나무꽃추출물, 알진	닥터포헤어	건성
1	C%84%B8&t_number=17	00018603736ko.jpg?l=ko	}완화 샴푸 750ml+200ml	4.7	12456	트80,1,2-헥산다이올,향료	닥터포헤어	듬,지성,건성,민감성,탈모
1	C%84%B8&t_number=18	00020261006ko.jpg?l=ko	획 (지성두피/가려운두피)	4.8	3938	드록시시트로넬알, 시트랄	닥터그루트	지성,건성,민감성,탈모
1	C%84%B8&t_number=19	00020466919ko.jpg?l=ko	선디셔너 750mL (2종 택1)	4.8	398	/메칠이소치아졸리논(3:1)	팬틴	민감성,탈모
1	C%84%B8&t_number=20	00020579504ko.jpg?l=ko	더블 기획 (430ml+430ml)	4.9	3737	출물, 시트릭애씨드, 향료	아도르	지성,건성
1	C%84%B8&t_number=21	00020621601ko.jpg?l=ko	드라이 샴푸 파우더 10g	4.7	314	<b>마연적으로 발견되는 성분</b>	아로마티카	지성,건성,민감성,탈모

# 2. 서비스 소개 2) 제품 추천 방식

## c. 두피 증상별 타입 매핑

두피증상	건성두피	지성두피	지루성두피	비듬성두피	탈모성두피	양호 두피
각질	<b>∨</b> ⊠					
피지		<b>∨</b> ⊠	<b>∨</b> ⊠			
모낭사이홍반		<b>∨</b> ⊠				
모낭홍반농포			<b>∨</b> ⊠			
비듬	<b>∨</b> ⊠			<b>∨</b> ⊠		
탈모					✓⊠	
양호						<b>∨</b> ⊠

## 2. 서비스 소개

### 2) 제품 추천 방식

d. 두피 증상별 타입 매핑 (예외 Case)

실제 모델을 돌려보니 앞선 두피 증상 이외의 경우 존재

- 1. 6개 증상 중에 하나만 높은 중증도를 가질 때
  - 각질 -> 건성
  - 피지 -> 지성
  - 모낭사이홍반 -> 설문 결과 반영 -> 상반될 시 모델링 결과 우선
  - 모낭홍반농포 -> 지루성두피
  - 비듬 -> 비듬성두피
  - 탈모 -> 탈모성두피
- 1. 6개 증상 모두 동일한 중증도(ex. 모두 경증)일 때
  - 사용자 설문 결과를 반영하여 지성, 건성 중에 하나로 결정
- 1. 가장 높은 중증도가 2개 또는 그 이상이면서 두피 증상 표 이외의 경우일 때
  - 모든 두피 증상과 관련된 제품 모두 추천

# **CONTENTS**

01

프로젝트 소개 <sup>목적</sup> 02

**서비스 소개** <sup>소개</sup>

제품추천 방식

03

데이터 소개

Al Hub

올리브영

04

모델링

제한사형

개선 과정

결과

05

하계

## a. 증상별 이미지 수

모델	두피유형	구분	양호(class1)	경증(class2)	중등도(class3)	중증(class4)	합	À
model1	모낭사이홍반	train	534	29,960	12,957	4,275	47,726	61,361
		val	152	8,560	3,702	1,221	13,635	
model2	모낭홍반농포	train	534	2,126	758	332	3,750	4,820
		val	152	607	216	95	1,070	
model3	미세각질	train	534	4,435	5,486	2,284	12,742	16,381
	증상의 유형	val	152	1,267 <b>今</b>	상별 증도	652	3,639	
model4	비듬	train	534	16,560	9,523	2,256	28,873	37,121
		val	152	4,731	2,721	644	8,248	
model5	탈모	train	534	13,346	3,797	836	18,513	23,801
		val	152	3,813	1,084	239	5,288	
model6	피지과다	train	534	28,061	24,485	3,746	56,826	73,062
		val	152	8,018	6,996	1,070	16,236	

## b. 증상별 두피 이미지













# c. 어노테이션 데이터(json)

항목	설명
Image_id	이미지 아이디
Image_file_name	파일명
value_1	미세각질(0, 1, 2, 3 중 1개 값)
value_2	피지과다(0, 1, 2, 3 중 1개 값)
value_3	모낭사이홍반(0, 1, 2, 3 중 1개 값)
value_4	모낭홍반농포(0, 1, 2, 3 중 1개 값)
value_5	비듬(0, 1, 2, 3 중 1개 값)
value_6	탈모(0, 1, 2, 3 중 1개 값)

증상별 각 4개의 Class = 각 증상의 중증도

# d. 메타 데이터 파일 구조(json)

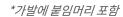
항목	설명
gender	성별
age	연령대
location	정수리 : TH, 좌측두 : LH 우측두 : RH, 후두부 : BH
question1(answer1)	샴푸사용 빈도(1일 1회, 1일 2회, 2일 1회)
question2(answer2)	펌주기 (하지않음, 1~3회/연, 4~6회/연, 7회 이상/연)
question3(answer3)	염색주기(자가 염색 포함) (하지않음, 1~3회/연, 4~6회/연, 7회 이상/연)

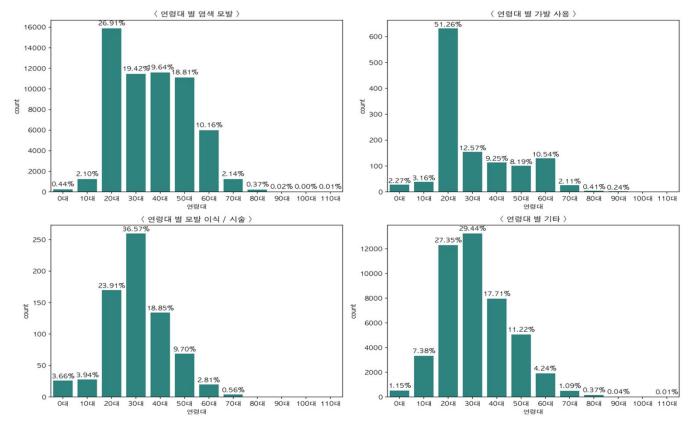
항목	설명
question4(answer4)	현재 모발 상태(염색모발, 기타, 가발사용(붙임머리 포 함), 모발이식/시술)
question5(answer5)	현재 사용하고 있는 두피모발용 제품 (샴푸, 트리트먼트, 헤어에센스, 린스, 헤어스타일링제, 두피스케일링제, 두피세럼)
question6(answer6)	맞춤 두피케어 제품 사용을 희망(선호)하시나요? (예, 아 니오)
question7(answer7)	샴푸 구매시 중요시 고려하는 부분은 무엇인가요? (세정 력, 두피자극, 머리결, 향, 헹굼후느낌, 가격)

### 1) Al Hub

### e. 메타 데이터 분석



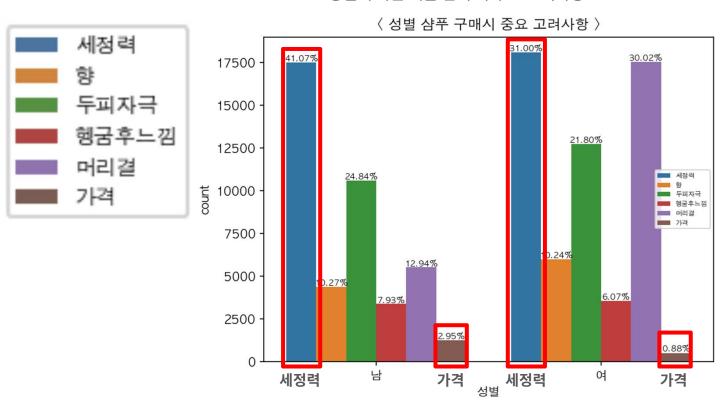




### 1) Al Hub

e. 메타 데이터 분석

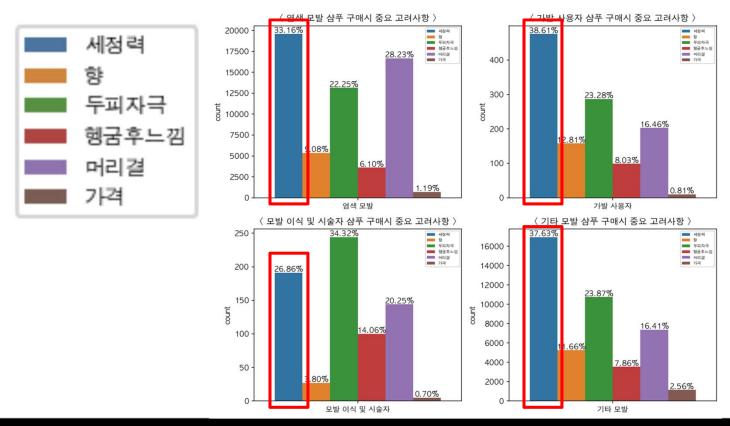




### 1) Al Hub

### e. 메타 데이터 분석





### 1) Al Hub

### e. 메타 데이터 분석

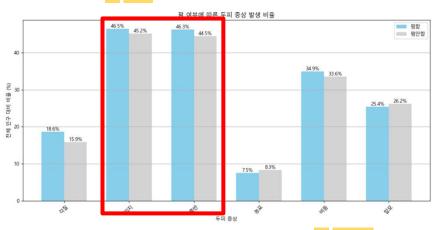
- 모든 증상별 중증도 폴더에 있는 이미지 파일을 겹치지 않게 목록 추출
- metadata 파일에 증상정보(value\_1~6) 연결 동일한 사람이지만 사진 부위에 따른 증상이 다른 경우 존재
- 이 경우 증상이 하나라도 있으면 증상이 있다고 판단하여 동일인 이미지 중복 제거

### 염색 여부에 따른 두피 증상 발생 비율



염색 여부에 따른 두피 증상 발생 비율 <mark>약</mark> 31%↑

### 펌 여부에 따른 두피 증상 발생 비율



펌 여부에 따른 두피 증상 발생 비율 약 2%↑

## 3. 데이터 소개 2) <sup>올리브영</sup>

a. 크롤링(데이터 수)



단위: (건)



# 3. 데이터 소개 2) <sup>올리브영</sup>

## b. 크롤링(내용)

컬럼	설명
page	제품 페이지
product_link	제품 링크
img_link	이미지 주소
brand_name	브랜드 이름
product_name	제품 이름
star	<mark>별점</mark>
review_count	리뷰 <mark>개수</mark>
ingrdient	<mark>성분</mark>

- 각 증상별 <mark>유효 성분 함유된</mark> 제품</mark> 크롤링
- 제품 추천 시, **별점**과 **리뷰 개수**를 동시에 고려하여 <mark>노출 순서 결정에</mark> 반영

page	product_lirimg_link product_name star		review_colingredient	brand_name
	1 https://ww.https://imi:어노브 실크 오일 헤어에센스 70ml 2종 택1 (텐더블룸/웜	4.8	19719 [텐더볼륨]	어노브
	1 https://ww.https://im:[단독기획]헤어플러스 단백질 본드 앰플 에센스 더블 기획	4.7	14752 정제수, 피마자씨오일, 세테아릴알코올, 다이프로필렌글라이콜, 아모다이메티콘, 스테아라미도프로필다이	메틸이헤어플러스
	1 https://ww.https://imi[대용량] 모로칸오일 헤어 트리트먼트 100ml(펌프포함) 탁	4.9	513 사이클로매치콘, 디메치콘, 아르간트리커넽오임, 향료, 아마인추출물, 적색225호, 황색204	모로칸오일
	1 https://ww.https://imi:미쟝센 컬링에센스 2X 볼륨컬/내추럴컬 230ml 택1	4.7	21530 정제수, 다이메티콘, 사이클로펜타실록세인, 글리세린, 다이프로필렌글라이콜, 브이피/메타크릴아마이드/	비닐이미쟝센
	1 https://ww.https://imi.바이브랩 리바이브 테라피 헤어 스칼프 앤 브로우 앰플	4.7	2266 정제수, 에탄율, 부틸렌글라이콜, 메틸프로판다이올, 글리세린, 1,2-헥산다이올, 피피지-13-데실테트라데서	스-24 바이브랩
	1 https://ww.https://imi[NEW] 미쟝센 퍼팩트 세럼 80ml [오리지널/스타일링/워티	4.8	47692 사이클로펜타실록세인, 다이메티콘을, 다이실록세인, C12-15알킬벤조에이트, 향료, 아르간커넬오일, 호호	바씨오미쟝센
	1 https://ww.https://imi[한정기획] 실크테라피 에센스 145KIT(130+15) 3중 택1 (	4.8	1961 사이클로펜타실록세인, 다이메티콘을, 변성알코올, 하이드롤라이즈드실크에틸에스터(5,000 ppm), C12-1:	5알킬빈실크테라피
	1 https://ww.https://imi[엔하이픈 성훈 PICK/체리우드 추가] 로레알파리 헤어오(	4.8	23798 [리치브라운오일]	로레알
	1 https://ww.https://imi[1&1더블] 어노브 실크 오일 헤어 에센스 더블 기획세트	4.8	19719 [텐더블룸]	어노브
	1 https://ww.https://imi[경민 PICK] 커리쉴 실키 오일 세럼 70ml	4.9	2710 사이클로펜타실록세인, 사이클로헥사실록세인, 다이실록세인, 다이메티콘을, 다이메티콘, 아이소프로필미	리스타커리쉴
	1 https://ww.https://im:미쟝센 퍼펙트 세럼 80ml 2개 (30ml 증정) [오리지널/스트	4.8	12705 사이클로펜타실록세인, 다이메티콘을, 다이실록세인, C12-15알킬벤조에이트, 향료, 아르간커넬오일, 호호	바씨오미쟝센
	1 https://ww.https://imi[대용량][NEW] 미쟝센 퍼펙트 오리지널 세럼 200ml	4.8	17478 사이클로펜타실록세인, 다이메티콘을, 다이실록세인, C12-15알킬벤조에이트, 향료, 아르간커넬오일, 호호	바씨오미쟝센
	1 https://ww.https://imi[쥴리 PICK] 피노 프리미엄 터치 헤어 오일 70ml 더블기호	4.8	817 하이드로제네이티드폴리아이소부텐, C13-14아이소알케인,다이메티콘율,아이소프로필미리스테이트,아이	소도데 피노
	1 https://ww.https://ima모로칸오일 헤어 트리트먼트 50ml (펌프 포함)	4.8	7090 사이클로메치콘, 디메치콘, 아르간트리커넬오일, 향료, 아마인추출물, 적색225호, 황색204호	모로칸오일
	1 https://ww.https://imi:어노브 볼륨 업 컬링 에센스 147ml (수분 컬크림)	4.7	6053 정제수, 다이메티콘, 다이프로필렌글라이콜, 글리세린, 브이피/브이에이코폴리머, 폴리아크릴아마이드, 프	I브이프어노브
	1 https://ww.https://imi[카즈하 PICK] 아도르 퍼퓸 헤어 오일 80ml 4중 택 1 (기회	4.9	6329 사이클로펜타실록세인, 다이메티콘, C12-15알킬벤조에이트, 다이메티콘을, 사이클로텍사실록세인, 아보키	가도오{아도르
	1 https://ww.https://imc클로란 퀴닌 에델바이스 두피 세럼 100ml(두피활력&모빌	4.7	1664 정제수, 에탄율, 글리세린, 키나나무껍질추출물(0.0485%), 에델바이스꽃/잎추출물(0.0375%), 망가니즈피세	시에이, 클로란
	1 https://ww.https://imi:모로칸오일 스무딩 로션 75ml	4.7	880 정제수, 사이클로메티콘, 다이메티콘, 세테아릴알코올, 글리세린, 향료, 베헨트라이모늄클로라이드, 폴리쿠	터늄- 모로칸오일
	1 https://ww.https://imi[쥴리 PICK] 피노 프리미엄 터치 헤어 오일 70ml	4.8	817 하이드로제네이티드폴리아이소부텐, C13-14아이소알케인,다이메티콘율,아이소프로필미리스테이트,아이	소도데피노
	1 https://ww.https://imi:모로칸오일 헤어 트리트먼트 라이트 50ml (펌프 포함)	4.8	1684 사이클로메치콘, 디메치콘, 아르간트리커넬오일, 향료, 아마인추출물, 적색225호, 황색204호	모로칸오일
	1 https://ww.https://imi[스마트톡 증정] 아도르 퍼퓸 헤어 오일 80ml 한정기획 2	4.9	6691 [히노키] 사이클로펜타실록세인, 다이메티콘, 다이메티콘올, C12-15알킬벤조에이트, 사이클로헥사실록세	인, 스 아도르
	1 https://ww.https://imi:그로우어스 데미지 테라피 두피 스케일링 앰플 50ml	4.6	144 정제수, 변성알코올, 프로판다이올, 글리세린, 슐베스-60테트라올리에이트, 1,2-헥산다이올, 잔탄검, 멘톨,	살리실그로우어스
	1 https://ww.https://ima달리프 글램 트루에센스 오리지널 더블 기획 (100ml+100	4.7	382 ■ 달리프글램 트루에센스 오리지널 100ml	달리프
	1 https://ww.https://im:롱테이크 헤어오일 50ml(단품/기획) 택 1	4.8	3109 사이클로펜타실록세인, 다이메티콘을, 다이실록세인, C12-15알킬벤조에이트, 코코넛야자오일, 향료(블랙	티앤피 롱테이크
	2 https://ww.https://imi[이보영 PICK] AZH 데일리 러스터 헤어오일_헤어 디자이	4.8	44 부틸렌글라이콜, 에탄올, 프로필렌글라이콜, 미리스틸알코올, 에톡시다이글라이콜, 아이소펜틸다이올, 정	제수, ∠AZH
	2 https://ww.https://imi[이보영 PICK] AZH 실키 에센스 헤어 오일_헤어 디자이니	4.9	26 사이클로펜타실록세인, 사이클로헥사실록세인, 아이소 프로필미리스테이트, 다이메티콘, 아이소도데케인	, 스쿠 AZH
	2 https://ww.https://ima박준뷰티랩 LPP 단백질 헤어오일 105ml (단백질 헤어 에	4.7	1702 사이클로펜타실록세인, 다이메티콘, 에탄올, 사이클로헥사실록세인, 트라이실록세인, 페닐트라이메	박준뷰티랩
	2 https://ww.https://im/모로칸오일 인텐스 컬 크림 75ml	4.7	2084 정제수, 미네탈오일, 세틸알코올, 폴리쿼터늄-7, 귤리세릴스테아레이트, 다이메티콘, 페트롤라돔, 아르간카	네널오일모로칸오일
	2 https://ww.https://ima롱테이크 샌달우드/블랙티앤피그 인텐시브 헤어오일 40r	4	1 다이메티콘, 세틸에틸렉사노에이트, 다이카프릴릴에터, C12-15알킬벤조에이트, 다이메티콘을, 아미노프트	르필다(롱테이크
	2 https://ww.https://imi[벨 PICK/NEW] 피노 프리미엄 터치 헤어 오일 세럼 에어	4.7	154 아이소도데케인,하이드로제네이티드폴리아이소부텐,C13-14아이소알케인,다이메티콘,다이메티콘율,아이	소프 피노
	2 https://ww.https://ima바이브랩 리바이브 테라피 헤어 스칼프 인핸싱 품	4.7	1003 편백수(51.6%),에탄율,정제수,부틸렌글라이콜,1,2-핵산다이올,카페인,C12-14알케스-12,폴리글리세릴-4카3	뜨레이 바이브랩

# **CONTENTS**

01

<mark>프로젝트 소</mark>개 <sup>목적</sup> 02

**서비스 소개** <sup>소개</sup>

제품추천 방식

03

데이터 소개 Al Hub

올리브영

04

모델링

제한사항

개선 과정

결과

05

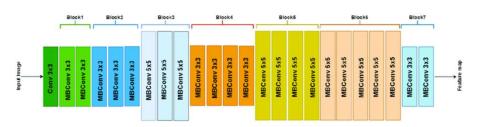
하계

# 4. 모델링 1) 어려웠던 점

	Problem	Reason	Solution
1차	- 전반적인 Accuracy 정체	- 데이터 불균형 - Pre-trained 모델링 시도 부족	- 다양한 Pre-trained 모델 학습 - focal loss 및 CosineAnnealingLR 스케줄러 적용
2차	- 전반적인 Accuracy 정체	- 라벨1과 라벨2 판별 한계	- 모델의 오분류 데이터 확인을 통한 라벨 병합 - 증상별 특징을 고려한 Gray scale 적용
3차	- 데이터 불균형 심화 - 피지과다 증상유무/정도 판단 제한	- 라벨(1과 2) 병합 - 증상 간 유사 이미지로 학습/판별 한계 (피지과다)	- 이미지수에 따른 가중치 적용(탈모) - 피지과다-미세각질 증상 간 높은 관련성 활용

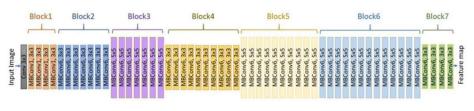
### 1) 어려웠던 점

- 기존 데이터셋 제공처의 모델 학습에 사용된 모델은 efficientNet b7으로 우리가 사용하기에 너무 큰 모델
  - -> 높은 비용 필요, 많은 학습 시간 소요
- 이미지 육안상으로 증상 구분 힘듦
- efficient b1 모델 및 이외 모델로 학습했으나 정확도가 70%대로 높이기 힘듦
  - -> 이외 사용 모델 : resnet, densenet, xception, efficientnet-b2(small), VIT(small)



### efficientNet b1

이미지 출처 : https://www.researchgate.net/figure/Architecture-of-EfficientNetB1-with-MBConv-as-basic-building-blocks-The-overall\_fig2\_351998710



### efficientNet b7

이미지 출처 : https://www.researchgate.net/figure/Architecture-of-EfficientNetB7-with-MBConv-as-basic-building-blocks-The-overall\_fig3\_343271189

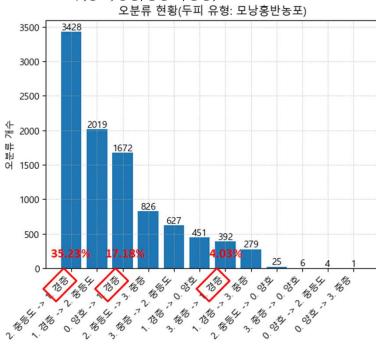
### 2) 개선 과정 (기존 VS 개선)

- efficientNet b7 보다 가벼운 모델 기준으로 데이터 증강, 파인튜닝, epoch, Learning Rate 등 변화시켜가면서 학습
- 1, 2% 정도의 차이는 있었지만 유의미한 개선은 보이지 않는 등 대부분의 테스트에서 <mark>70% 초반의</mark> <mark>정확도</mark>를 보임

-> 모델에 따른 정확도 차이가 크지 않는다는 점 확인

모낭사이홍반, 피지과다	특이사항	모낭홍반농포, 비듬	특이사항	탈모 / 피지과다 / 미세각질	특이사항	피지과다 / 미세각질	특이사항
1. efficientnet-b0(모낭, 78.2%, 5, 10/15)	1. A100 GPU, 23분/epoch, 59분 크기: 600	1. efficientnet-b1 (70.1%)	1. 에포크 5, 1당 약 5분, 사이즈 224 & Ir 1e-4	1: efficientnet-b1 (76.17%)	1: epoch-5, 25분/epoch, focal loss 적용	1.efficientnet-b1 (65.4%)	1 - epoch : 5, 개당 70분
2. efficientnet-b1(모낭, 78.5%, 5, 10/15)	2. A100 GPU, 30분/epoch, 1시간 17분 크기: 600	2. Resnet 50 (73.1%)	2. 에포크 15, 1당 약 1분, 사이즈 220 & Ir 1e-4	2: efficientnet-b0 (72.67%)	2: epoch-5, 13분/epoch	2.efficientnet-b1 (66%)	2 - epoch : 5, 개당 30분
3. efficientnet-b1(ℤ ス , 66.1%, 5, 10/15)	3. L4, 22분/epoch, 56분, 크기 : 240	3. efficientnet-b0 (73.8%)	3. 에포크 5, 1당 약 1분, 사이즈 224 & lr 1e-4	3: efficientnet-b1 (73.7%)	3: b1, epoch-6, 15분/epoch, 가중치?	3. efficientnet-b1 (65.2%)	3 - epoch : 5, 개당 30분 학습률, 데이터 증강 변경
4. resnet50 + ft-s3(모낭, 73.9%, 5, 10/16)	4. L4, 5분/epoch, 12분 26초, 크기 : 240	4. efficientnet-b0 (69.6%)	4. 에포크 10, 1당 약 4분, 사이즈 600 & Ir 5e-4	4: efficientnet-b1 (73.7%)	4: epoch-10, 4분/epoch, 가중치 + 다른방법 적용 -> 별차이없음	4. efficientnet-b1 (68%)	4 - epoch : 10, 개당 10분 옵티마이저 AdamW 변경, 배치사이즈 12
5. resnet50 + ft-s2(모낭, 77.9%, 5, 10/16)	5. L4, 7분/epoch, 16분 48초, 크기 : 240	5. efficientnet-b0 (64.8%)	5. 에포크 5, 1당 약 4분, 사이즈 600 & Ir 1e-4	5: efficientnet-b1 (77.0%)	5: epoch-10, 4분/epoch, 이미지 사이즈 비율 유지한채로 줄이기, focal loss 적용	5. efficientnet-b1 (62%)	5 - epoch : 10, 개당 20분 파인튜닝 적용 (마지막 레이어만 확습)
6. efficientnet-b1+ft-s3(모낭, 68.3%, 5, 10/17)	6. L4, 9분/epoch, 23분 38초, 크기 : 240	6. efficientnet-b0 (70.7%)	6. 에포크 5, 사이즈 224, augment color 추가	6: efficientnet-b1 (79.0%)	6: epoch-15, 위와 동일, + CosineAnnealingLR 스케쥴러 사용	6. efficientnet-b1 (66.4%)	6 - epoch : 10, 개당 15분 클래스 가중치 적용
7. efficientnet-b1+ft-s3(모낭, 71.7%, 5, 10/17)	7. L4, 21분/epoch, 50분 38초, 크기 : 600, 불일치 csv 파일저장	7. efficientnet-b0 (71.5%)	7. 에포크 5, 사이즈 224, optimizer SGD로 변경			7. resnet50 (58%)	7 - epoch : 5, 개당 7분
8. efficientnet-b1+ft-s3(모낭, 59.8%, 5, 10/17)	8. L4, 9분/epoch, 23분 17초, 크기 : 코드, 불일치 csv 파일저장	8. efficientnet-b0 (71.6%)	8. 에포크 5, 사이즈 224, Ir 5e-4			8. densenet121 (61.3%)	8 - epoch : 5, 개당 8분
9. resnet50(모낭, 78%, 5, 10/18)	9. L4, 10분/epoch, 26분 31초, 크기 : 224	9. efficientnet-b0 (70.8%)	9. 에포크 10, 사이즈 224, Ir 5e-4			9. efficientnet-b1 (65%)	9. epoch : 10, 개당 5분 FP16 혼합 정밀도 확습 적용

- train / validation Dataset에 대한 <mark>오분류 이미지 리스트업</mark>
- 두피유형: 모낭홍반농포에서 <mark>경증으로</mark> 오분류 <mark>비율</mark> 56.44%
- Label 4개 (양호/경증/중등도/중증)





0013\_A2LEBJJDE000600\_ 1604372042927\_5\_RH 모낭홍반농포(1\_경증)

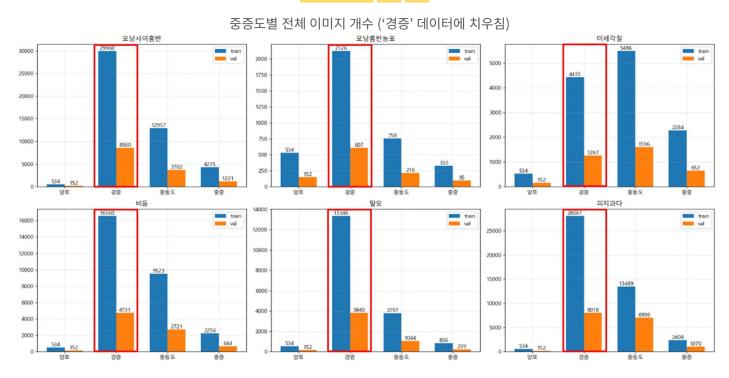


0013\_A2LEBJJDE000600\_ 1605534137037\_6\_BH 모낭홍반농포(2\_중등도)



1080\_A2LEBJJDE00092R\_ 1603792861159\_5\_RH 모낭홍반농포(3\_중증)

- 기존 4개 Label에서 증상에 따라 이미지 개수 차이가 심한 문제 발견
- → 다수의 증상 이미지에서 경증 이미지 수가 절반이상 차지 → <mark>데이터불균형 문제 발생</mark>



- 높은 정확도를 보여줬던 efficientNet b1 모델을 기본으로 하여 데이터 불균형 문제 해결 방안 모색
  - **Focal Loss** : Focal Loss는 <mark>쉬운 예제(즉, 모델이 이미 잘 맞추고 있는 예제)의 손실을 줄이고, 어려운 예제(잘못 예측하고 있는 예제)의 손실에 더 큰 가중치 부여</mark> 즉, 모델이 잘못 예측하기 쉬운 클래스나 상대적으로 갯수가 적은 클래스에 더 집중해서 학습하게 되어 불균형한 데이터셋에서 좋은 성능을 낼 수 있게 함
  - **CosineAnnealingLR 스케줄러** : 학습률을 코사인 곡선 형태로 변화시키는 스케줄러로, 설정한 `T\_max`에 따라 학습률이 주기적으로 최소값 (`eta\_min`)에 도달한 후 다시 증가. <mark>주로 학습률 점진적으로 감소, 과적합을 방지</mark>하며 모델의 성능을 최적화하는 데 유용
    - -> 탈모 모델 기준 79% 정확도 달성
- 하지만 다른 증상 모델의 경우, 이전보다는 약간 개선이 되긴 했으나 역시 큰 차이는 없음

### 2) 개선 과정 (기존 VS 개선)

- 오분류된 목록을 확인해보니 '라벨 1(경증)'과 '라벨 2(중등도)' 판별을 어려워하는 점 발견 → 육안으로도 판별 한계
- 기존 클래스 4개에서 <mark>클래스</mark> <mark>3개로 변경</mark>(모델별 오답 비율 확인 후 모두 기존 라벨 1, 2 를 모두 1로 병합)



클래스 병합 이후 오분류 현황

- 클래스 3개
- -> 양호를 경증 or 중등도로 오분류 케이스 약 81% 감소
- -> 중증을 경증 or 중등도로 오분류 케이스 약 59% 감소
- 클래스1(경증\_중등도)과 2를 오분류하는 숫자 늘었으나, 증상의 정도보다 유무를 정확히 판단하는 것이 더 중요하다고 판단.

- 피지과다 증상 보유 & 미세각질 양호인 경우 97%
- 카이제곱검정으로 두 증상 간의 유의미한 연관성이 있는지 검정
- 카이제곱 통계량이 무척 크고, p-value가 매우 작아 두 증상이 유의미한 연관성이 있다고 판단

```
[8] from scipy.stats import chi2_contingency
contingency_table = pd.crosstab(df['미세각질'] != 0, df['피지과다'] == 0)
chi2_stat, p_value, _, _ = chi2_contingency(contingency_table)
print(f"- 미세각질 중상이 있고, 피지과다는 양호 -\nchi2 통계량 : {chi2_stat}\np-value : {p_value}")

- 미세각집 중상이 있고, 피지과다는 양호 -
chi2 통계량 : 47650.266239731965
p-value : 0.0

from scipy.stats import chi2_contingency
contingency_table = pd.crosstab(df['미세각질'] == 0, df['피지과다'] != 0)
chi2_stat, p_value, _, _ = chi2_contingency(contingency_table)
print(f"- 미세각질이 양호, 피지과다가 중상이 있음 -\nchi2 통계량 : {chi2_stat}\np-value : {p_value}")

- 미세각질이 양호, 피지과다가 중상이 있음 -\nchi2 통계량 : {chi2_stat}\np-value : {p_value}")
```



### 2) 개선 과정 (기존 VS 개선)

- 오분류의 72%가 비듬 라벨 1, 2를 잘못 판별한 점 확인
  - → 직접 해당 이미지를 확인 후, 이미지에서 증상 유무에 따른 명암 차이 존재 확인
  - → 이미지에 그레이 스케일을 적용하여 명암 강조 후 학습하였으며, 정확도 개선 확인
- 비듬과 비슷한 이슈가 있는 두피유형에 적용 (증상 유무에 따른 명암)
  - 적용: 미세각질, 비듬, 피지과다
  - 비적용: 모낭홍반농포, 모낭사이홍반, 탈모
- 이전보다 정확도는 증가했지만 그럼에도 특정 라벨에서의 높은 오류율 존재
  - → ex. 탈모: 2 잘 못맞춤
  - → ex. 모낭사이홍반: 0 잘 못맞춤

7.949791

- 기존 모델에 추가적으로 <mark>클래스별 데이터수에 따른 가중치를 동적으로 조정</mark>하여 전 증상 모델 재학습 테스트
  - → 이전과 차이가 없는 모델(ex. 모낭사이홍반)은 그대로 사용

→ 개선이 된 모델은 해당 모델 사용(ex. 탈모) 이전 탈모 중증(2) 맞춘 확률

image 7.949791 top\_class 7.949791 top\_probability 7.949791 second\_class 7.949791 second\_probability 7.949791 top\_score\_range 7.949791

second\_score\_range

dtype: float64

Accuracy 약 49% 개선

image 56.485356
top\_class 56.485356
top\_probability 56.485356
second\_probability 56.485356
second\_probability 56.485356
second\_score\_range 56.485356
dtype: float64

이후 탈모 중증(2) 맞춘 확률

# 3) 모델링 결과

- 이미지비율크기조정, augment color, focal loss, CosineAnnealingLR 스케줄러 클래스1, 2 병합

모델	두피유형	모델	추가/변경옵션	accuracy
model1	모낭사이홍반	Efficientnet-b1	고정크기 : 600	93.2%
model2	모낭홍반농포	Efficientnet-b1	-	89.5%
model3	미세각질	Efficientnet-b1	그레이 스케일	84%
model4	비듬	Efficientnet-b1	그레이 스케일	92.7%
model5	탈모	Efficientnet-b1	동적 가중치 적용	93.5%
model6	피지과다	Efficientnet-b1	그레이 스케일	92.6%

# 서비스 시연

PC: https://9tcrmtebpkzuvxi8c6srx5.streamlit.app/

Mobile:

# **CONTENTS**

01

<mark>프로젝트 소</mark>개 <sup>목적</sup> 02

**서비스 소개** <sup>소개</sup>

제품추천 방식

03

데이터 소개 Al Hub

올리브영

04

포필딩 제하사한

....

11 11 11 11

결괴

05

한계

# 5. 한계

- 1. 웹에 모델을 실어야했기에 <mark>파라미터 수가 많은 모델 사용 제한</mark>
- 1. 육안으로도 <mark>구분하기 힘든 이미지들이 많아</mark> 여러 튜닝기법을 사용해도 정확도 향상이 어려웠음
- 1. 로컬 환경에서 모델 학습을 진행하기 어려워 코랩을 활용하였는데 기본 제공되는 GPU는 느리고 <mark>사용량의</mark> 한 계가 있어 <mark>추가 비용이 발생(약</mark> 20만 원)

# 감사합니다