

데이콘 HAI(하이)! – Hecho AI Challenge

차량 분류 모델 프로젝트

2025.06

서동현, 유동현, 김준희

배경 (Background)

자동차 산업의 디지털 전환이 가속화되면서 차종을 빠르고 정확하게 인식하는 기술의 중요성이 크게 증가
특히 중고차 거래, 차량 관리, 자동 주차·보안 시스템 등에서 정확한 차종 분류는 핵심 경쟁력으로 자리잡음

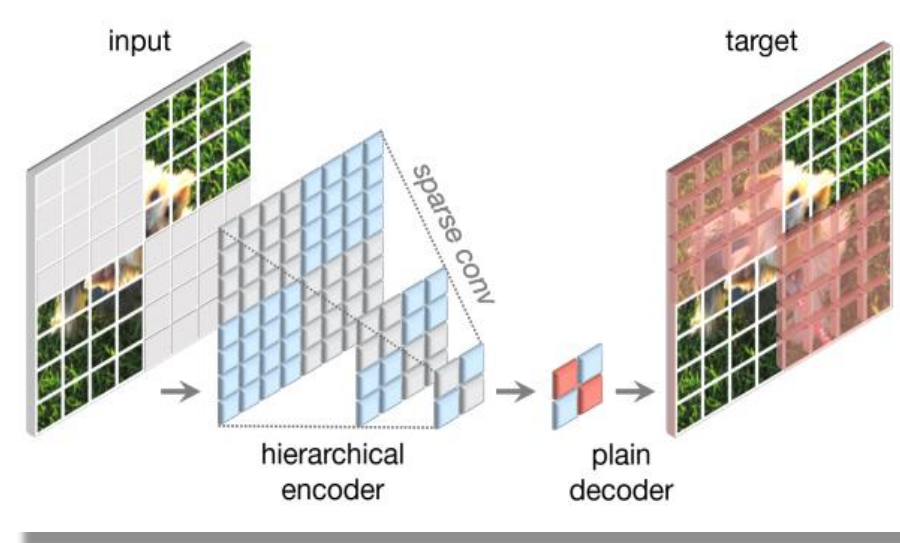
프로젝트 개요

데이터

모델링

향후 방향성

사용 분류 모델



ConvNeXt는 ResNet을 기반으로 Conv 블록을
현대화해 대규모 데이터 학습과 깊은 네트워크
설계를 지원하는 차세대 CNN 아키텍처

평가산식

$$LogLoss = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C y_{i,j} \log(p_{i,j})$$

- N : 전체 샘플 수 (이미지 개수)
- C : 클래스 수 (차량 종류 개수)
- $y[i][j]$: i번째 샘플의 정답 클래스가 j이면 1, 아니면 0 (one-hot 형식)
- $p[i][j]$: i번째 샘플에 대해 모델이 클래스 j라고 예측한 확률 (softmax 결과)
- 예측 확률과 실제 라벨 간의 교차 엔트로피 손실

데이터 소개 (Data Info.)

프로젝트 개요

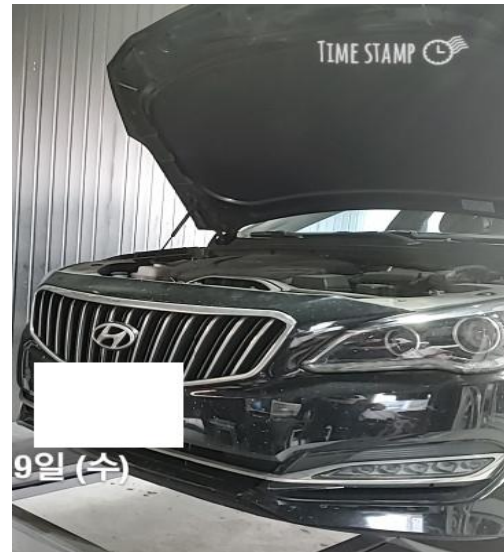
데이터

모델링

향후 방향성

Train Data set

✓ 396개의 각 Class 폴더 내 jpg 파일 (총 33,137장)



Test Data set

✓ 평가 데이터셋 (총 8,258장)



데이터 전처리 (Data Preprocessing)_1

Ver.1 각 클래스별 모델 학습에 노이즈라고 생각되는 이미지를 직접 눈으로 보고 제거

프로젝트 개요

데이터

모델링

향후 방향성

원본 train 데이터
33,137장



497장
노이즈선별



train 데이터
32,640장

전처리 대상 예시



차량이미지 훼손



과도한 비침



측면 내부 이미지



트렁크 열림



차량이미지 훼손

데이터 전처리 (Data Preprocessing)_2

Ver.2 차량사진이 온전히 보존되어 있지 않거나, test데이터에 해당 이미지와 유사한 유형이 존재 하지 않았을경우 제거

원본 train 데이터
33,137장



63장
노이즈선별

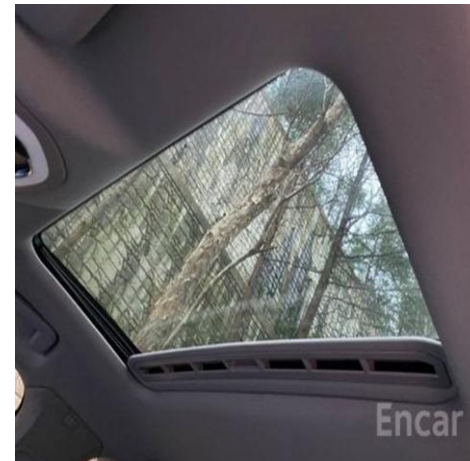


train 데이터
33,074장

최종 전처리 대상 예시



차량 내부 이미지



분류 불가 이미지



식별 불가 이미지

데이터 탐색 (EDA)_이미지 크기 분포 분석

프로젝트 개요

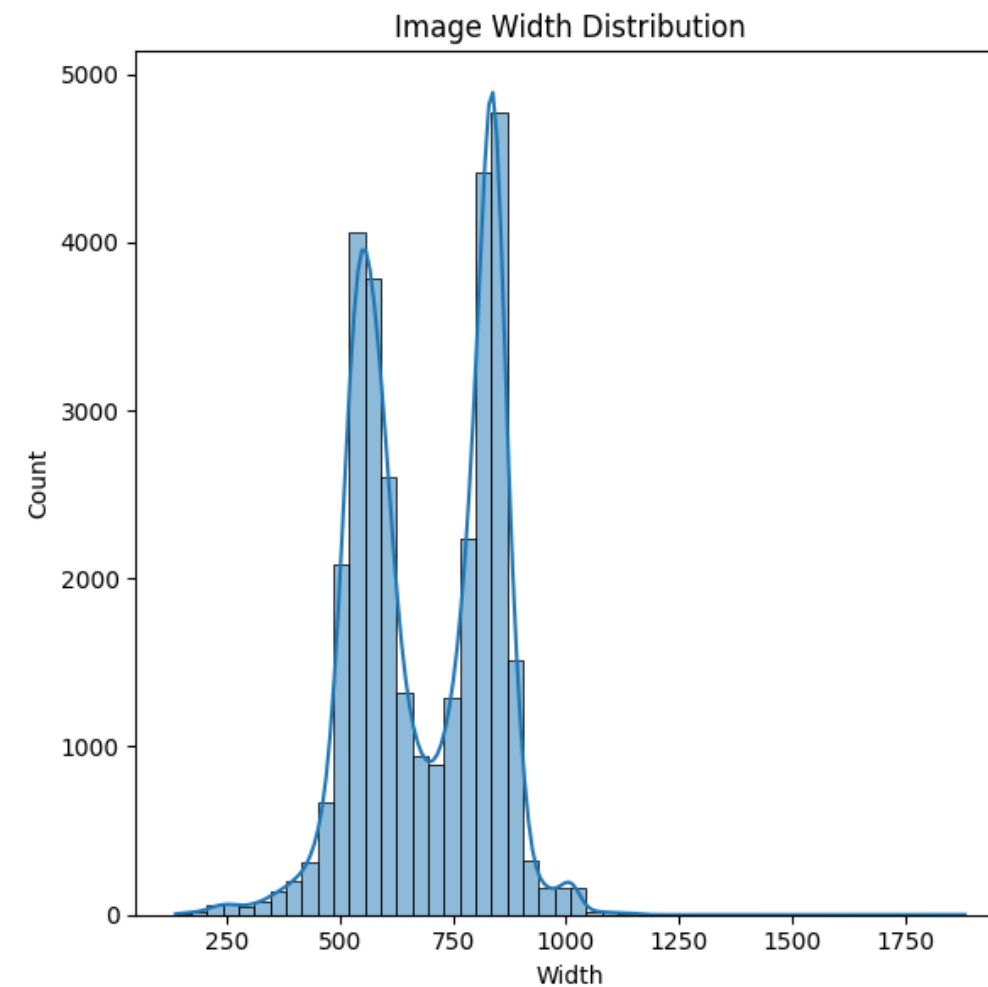
데이터

모델링

향후 방향성

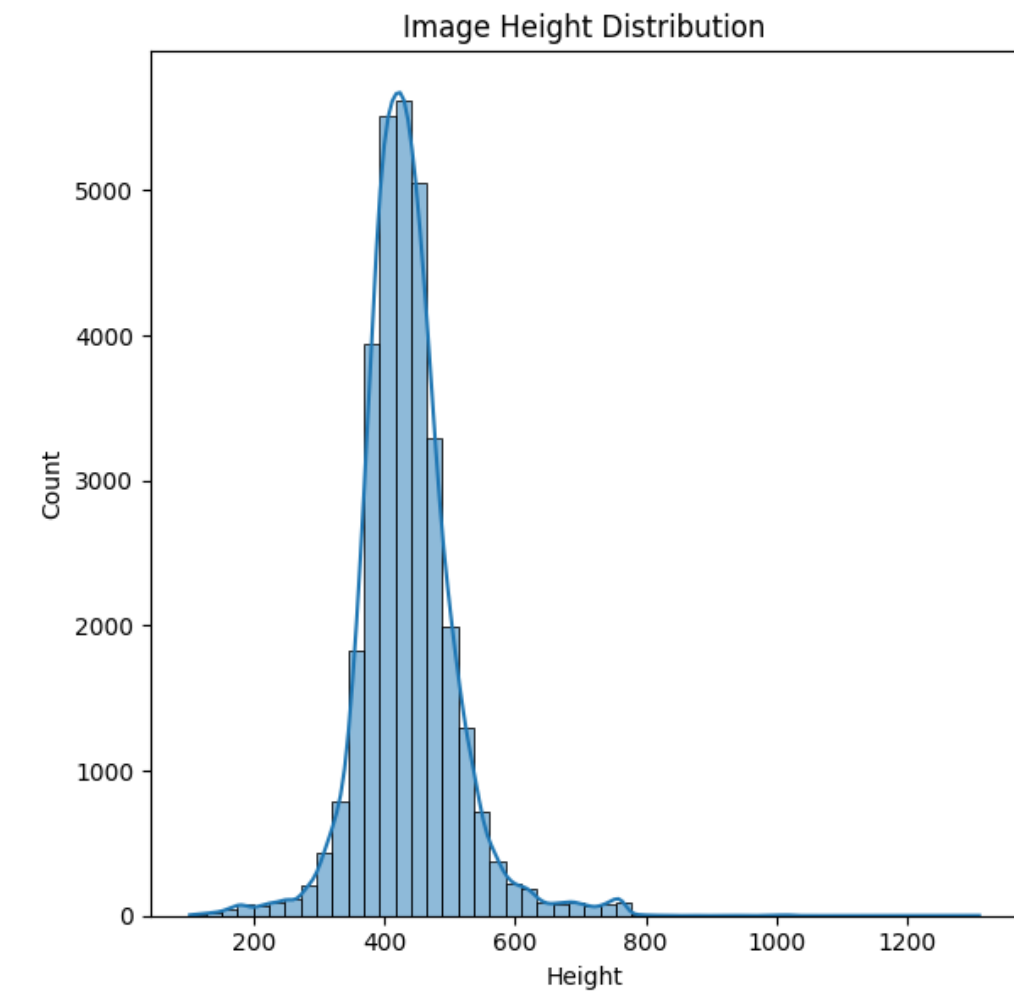
이미지 너비(Width) 분포

- ✓ 집중 구간: 500-900픽셀 (전체의 약 80%)
- ✓ 최빈값: 약 800픽셀 근처
- ✓ 분포 형태: 600픽셀과 800픽셀 근처에서 피크



이미지 높이(Height) 분포

- ✓ 집중 구간: 400-600픽셀 (전체의 약 85%)
- ✓ 최빈값: 약 450-500픽셀 근처
- ✓ 분포 형태: 대칭적 정규분포에 가까움



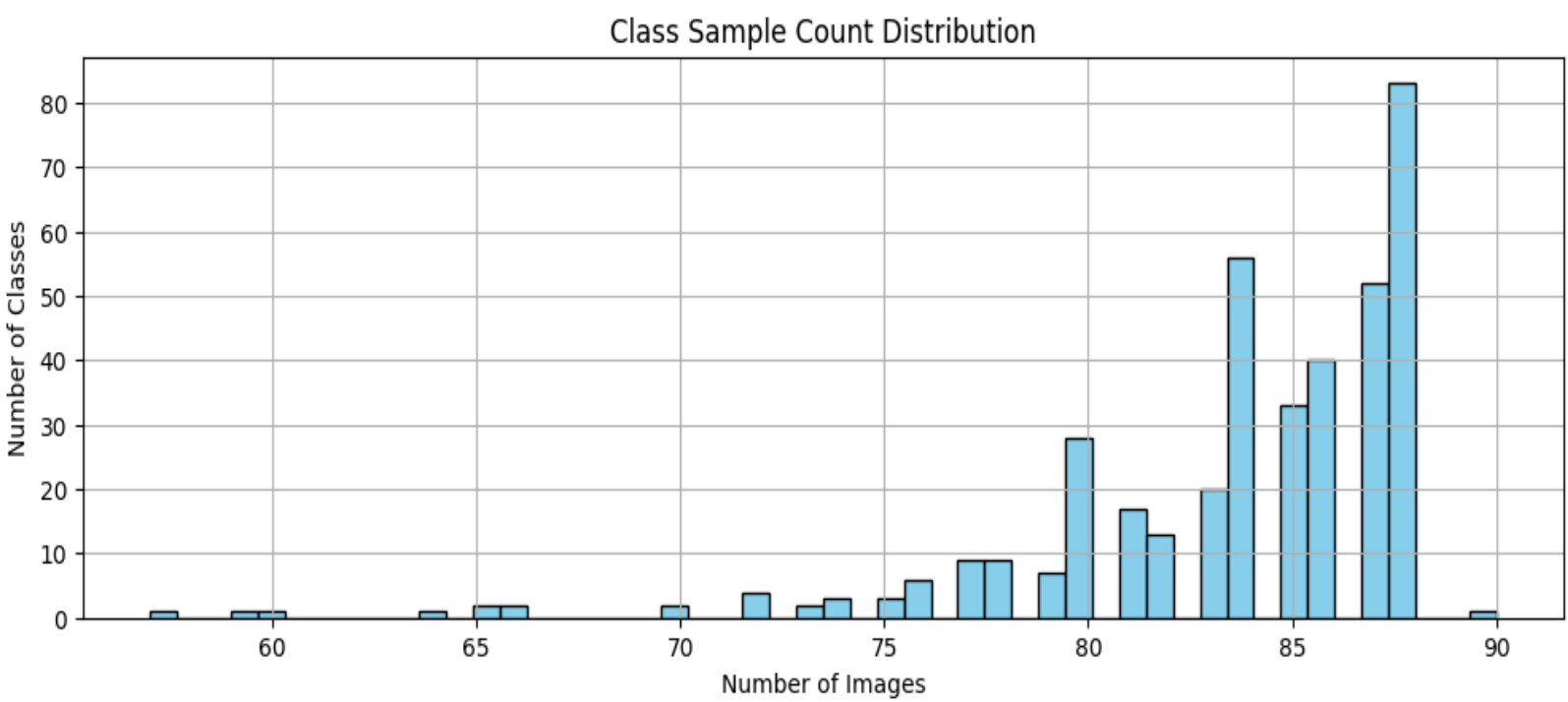
데이터 탐색 (EDA)_클래스별 이미지 수 분석

클래스 불균형

- ✓ 클래스당 이미지 수: 57개 ~ 88개 (약 31개 차이)
- ✓ 최빈 구간: 85~88개 이미지를 가진 클래스가 가장 많음
- ✓ 분포 형태: 우편향 분포(대부분 클래스가 많은 이미지 보유)

분포패턴

- ✓ 소수 클래스: 55~57개 이미지 (소수)
- ✓ 다수 클래스: 80~88개 이미지 (대다수)
- ✓ 집중도: 전체 클래스의 약 70%가 80개 이상의 이미지 보유



데이터 증강 (Data Augmentation)_1

프로젝트 개요

데이터

모델링

향후 방향성

AutoAugment

- ✓ Google에서 개발한 자동 증강 정책 탐색 기법
- ✓ 강화학습을 통해 데이터셋에 최적화된 증강 조합을 자동으로 찾음
- ✓ 회전, 색상 변경, 전단 변환 등 다양한 변환을 조합

AutoAugment



RandAugment



RandomAugment

- ✓ AutoAugment의 단순화 버전
- ✓ 무작위로 선택된 N개의 변환을 동일한 강도로 적용
- ✓ 계산 비용이 낮으면서도 효과적인 성능 향상

AutoAugment



RandAugment



데이터 증강 (Data Augmentation)_2

프로젝트 개요

데이터

모델링

향후 방향성

Mixup

- ✓ 두 이미지를 선형 결합하여 새로운 훈련 샘플 생성
- ✓ 라벨도 동일한 비율로 혼합하여 부드러운 경계 학습
- ✓ ALPHA (강도) : 0.3 ~ 0.5



CutMix

- ✓ 한 이미지의 일부를 다른 이미지로 교체
- ✓ Cutout과 Mixup의 장점을 결합한 기법
- ✓ ALPHA (강도) : 0.3 ~ 0.5



Mixup/Cutmix 기법이 적용되었을 때 **확률적으로 mixup/cutmix 중 하나를 적용**
Ex.) **Mixup/Cutmix**가 30%확률로 적용 → **Mixup** 이 60%확률로 적용 또는 **Cutmix** 적용

데이터 증강 (Data Augmentation)_3

프로젝트 개요



데이터

모델링

향후 방향성

Crop & Pad	<div>✓ 이미지를 자르고 패딩을 추가하는 기법</div> <div>✓ 객체의 위치 변화를 통한 다양성 확보 시도</div>	
RandomResizedCrop	<div>✓ 무작위 크기로 자르고 리사이즈</div> <div>✓ 객체의 스케일 변화에 대한 적응력 향상 시도</div>	
RandomShadow	<div>✓ 무작위 그림자 효과 생성</div> <div>✓ 실제 환경의 다양한 그림자 조건 시뮬레이션</div>	
RandomSunFlare	<div>✓ 태양 플레어 효과를 무작위로 추가</div> <div>✓ 조명 조건 변화에 대한 강건성 향상 목표</div>	

데이터 증강 (Data Augmentation)_4

Mosaic	<ul style="list-style-type: none">✓ 4개 이미지를 조합하여 하나의 훈련 이미지 생성✓ YOLO 계열에서 효과적이나 분류 태스크에서는 부적합		
RandomErasing	<ul style="list-style-type: none">✓ 이미지의 임의 영역을 사각형으로 마스킹✓ 모델이 객체의 일부만 보고도 분류할 수 있도록 학습✓ 과적합 방지 및 일반화 성능 향상		

부작용

- ✓ 과도한 변형 - 원본 이미지의 특성을 너무 많이 변화시킴
- ✓ 노이즈 증가 - 불필요한 시각적 노이즈로 인한 학습 방해
- ✓ 태스크 부적합 - 분류 문제에 적합하지 않은 기법들



선별적 증강 적용의 **중요성** 확인
모든 증강 기법이 **항상 효과적이지 않음**
데이터셋과 태스크 **특성에 맞는 기법 선택** 필수

실험 모델 목록 (Model Selection)

프로젝트 개요

데이터

모델링

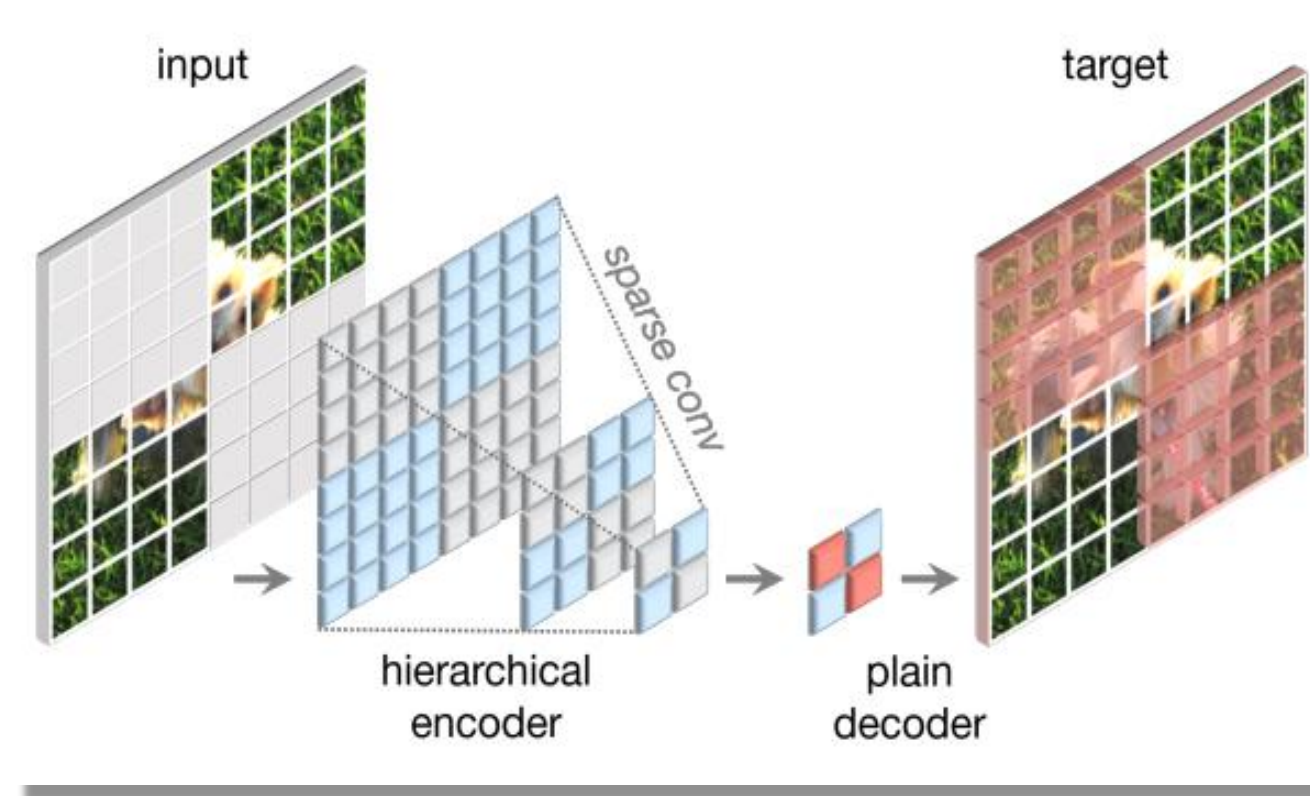
향후 방향성

모델	모델 상세 내용	LB 점수
Swin Transformer Microsoft의 계층적 Vision Transformer	Shifted Window 기반 효율적 Self-Attention 지역적 윈도우에서 전역적 정보로 점진적 확장 선형 복잡도로 고해상도 이미지 처리 가능	0.2259
EVA-02 BAAI의 대규모 Vision Transformer	Masked Image Modeling + Multi-scale Training 38M 이미지로 사전훈련된 강력한 표현 학습	0.2002
DeiT v3 Meta의 지식 증류 기반 ViT	Teacher-Student 증류 + CaiT 블록 구조 38M 이미지로 사전훈련된 강력한 표현 학습	0.1799
BEiT v2 Microsoft의 자기지도 학습 ViT	Vector-Quantized Knowledge Distillation DALL-E의 dVAE 토크나이저를 활용한 이미지 토크화 마스크된 패치를 시각적 토크으로 예측하는 사전훈련	0.2166
ConvNeXt v2 Meta에서 개발한 현대적 CNN 아키텍처	ResNet의 구조를 Transformer 설계 원칙으로 개선 Depthwise Convolution + LayerNorm + GELU 활성화 Global Response Normalization (GRN) 적용	<div>✓</div> <div>0.1718</div>

모델 선정 (Model Selection)

ConvNeXt v2

계층적 구조로 전체 형태부터 세부 형태까지 점진적 특징 추출이 가능하고 CNN의 지역적 패턴 학습 능력이 차량 도메인에 최적화 차량의 그릴, 헤드라이트, 로고 등 세밀한 특징을 효과적으로 포착할 수 있어 396개 세분화 클래스 구분에 적합 Base/Large 등 다양한 버전을 통한 성능-비용 트레이드오프 실험 가능 검증된 아키텍처로 안정적으로 학습 보장



실험, 학습시간, 모델 안정성이 다른 모델의 비해 월등하다 할 수는 없으나 각 모델의 실제 테스트 결과 성능면에서 ConvNeXt v2의 성능이 가장 높게 평가 되어 해당 모델을 선정

여러 기법들을 추가로 실험하여 성능 향상을 확보하려 함

추가 기법 (Additional Technique)

프로젝트 개요

데이터

모델링

향후 방향성

1. 최적화 기법 (Optimization Techniques)

SWA (Stochastic Weight Averaging)
여러 체크포인트의 가중치 평균화
손실 landscape의 flat minimum 탐색으로 일반화 성능 개선

EMA (Exponential Moving Average)
모델 파라미터의 지수 이동 평균 유지
훈련 과정의 노이즈 완화로 안정적 성능 향상

CV와 LB간 점수 격차 완화

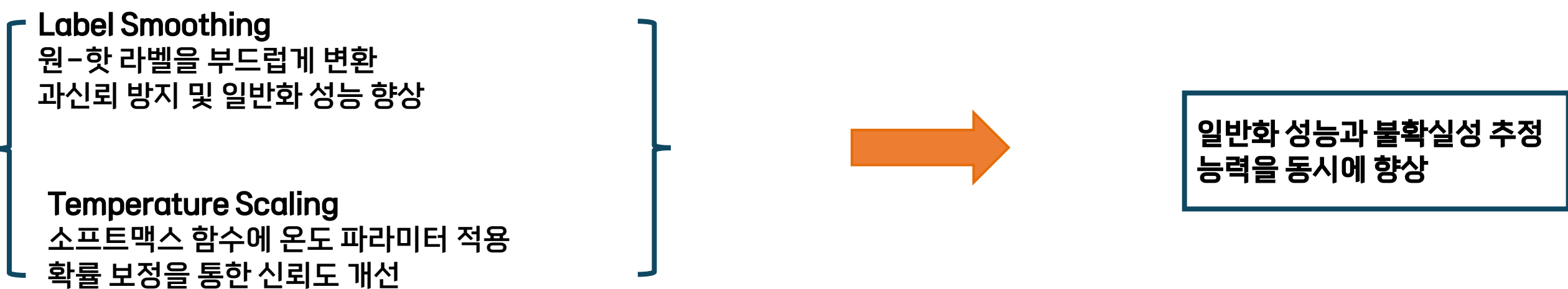
2. 학습률 스케줄링 비교 적용

	설명	점수 (Log Loss)
✓ Warmup+ CosineAnnealing	낮은 학습률로 시작하여 점진적 증가 코사인 함수 형태로 학습률 감소	0.1558
Reduce LROnPlateau	검증 지표 개선 정체 시 학습률 동적 감소 적응적 학습률 조정으로 최적화 효율성 향상	0.1575

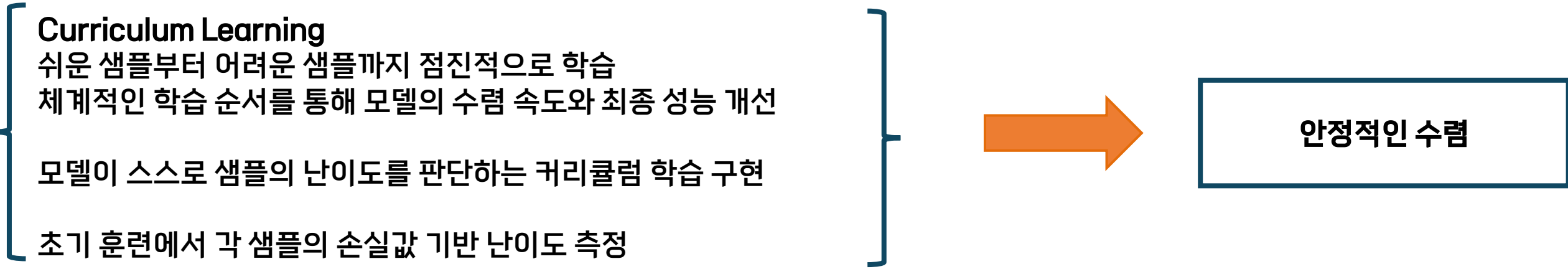
학습 및 점수 안정성 확보

추가 기법 (Additional Technique)

3. 정규화 및 보정 기법



4. 학습 전략(Data Scheduling)



최종 학습 모델 (Final Trained Model)

프로젝트 개요

데이터

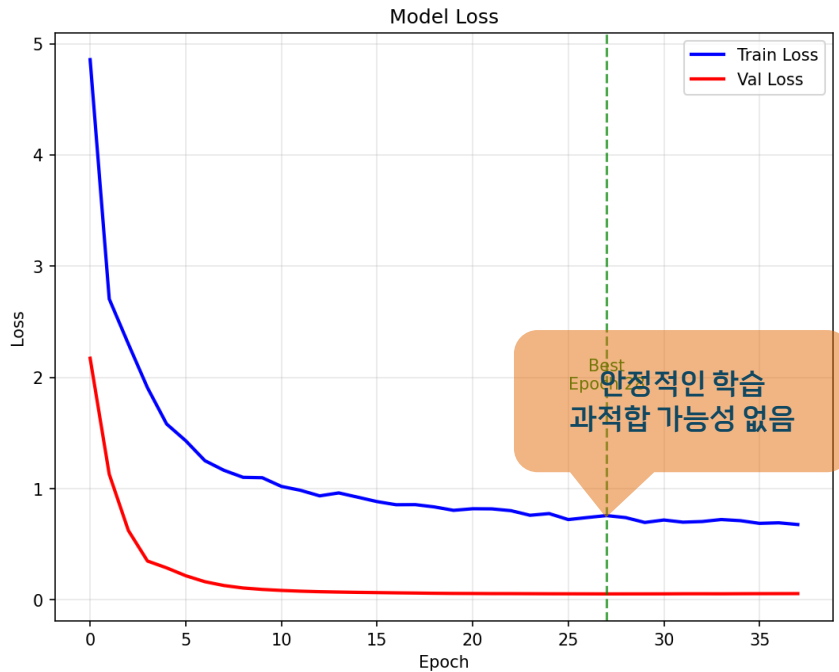
모델링

향후 방향성

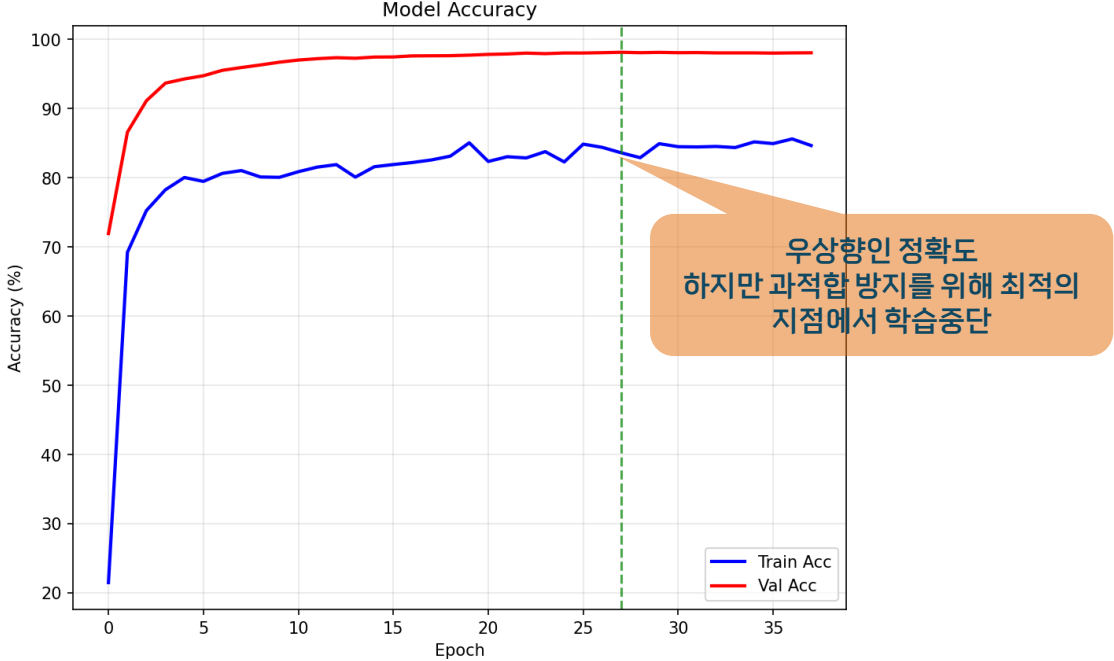
ConvNeXt v2 기반 모델 파라미터 정의

batch_size	배치 크기	64
learning_rate	학습률	5e-5
epochs	에포크 수	50
Image_size	이미지 크기	384/512
use_ema	EMA 사용 여부	True
ema_decay	EMA 감쇠율	0.999
ema_warmup_epochs	EMA 워밍업 에포크	5
mixup	Mixup 증강	True
cutmix	CutMix 증강	True
weight_decay	가중치 감쇠	0.05
label_smoothing	라벨 스무딩	0.05
early_stopping_patience	조기 종료 인내값	10

모델 Loss



모델 정확도



추론 및 앙상블 (Inference & Ensemble)

프로젝트 개요

데이터

모델링

향후 방향성









Test-Time Augmentation (TTA)

테스트 이미지에 다양한 변환(회전, 플립, 스케일링) 적용
각 변환된 이미지에 대한 예측 결과의 평균 계산
단일 이미지 대신 여러 변환 결과를 활용하여 예측 안정성 향상

Ensemble(Soft Voting 기반)

여러 모델의 클래스별 확률 분포를 평균화
각 모델의 예측 신뢰도를 종합적으로 활용
Hard Voting 대비 더 풍부한 정보 활용

Public LB : 0.11653 (LogLoss)
Private LB : 0.12452 (LogLoss) – 36위

30	월봐		0.11977	83
31	정연비		0.12078	43
32	오곡밥		0.12081	72
33	sseo		0.12136	72
34	oliviatoto		0.12288	5
35	비회원		0.12317	5
36	코딩대한민국		0.12452	108
37	부캠즈		0.12469	101

주차 차량 자동 분류 시스템

중고차량의 사진을 자동 라벨링 시스템(디텍션 시스템)만 구축되어 있다면 분류모델을 이용하여 차종을 정확히 분류할 수 있는 시스템을 구축
가능할 것으로 예상
이를 통해 중고차 정보를 연식, 차종에 대해 오기재를 방지하여, 구매자와 중고차 플랫폼의 피해를 최소화할 것을 기대

프로젝트 개요

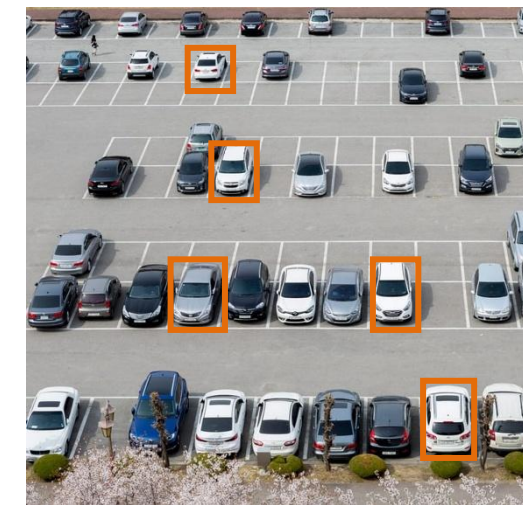
데이터

모델링

향후 방향성



디텍션 시스템



차량 라벨링



자동차량 분류

정확한
정보전달



감사합니다.

2025.06

서동현, 유동현, 김준희