

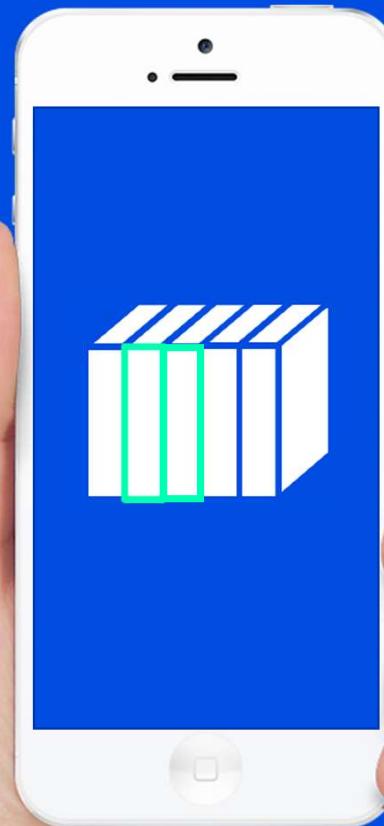


SeeBook-OCR

Object detection과 OCR을 활용한
여러 권의 책 정보 검출 및 검색 서비스

책봐조

박진우 | 김성연 | 나웅 | 장유진



목차

01

배경 및 필요성

02

개요 및 목표

03

데이터
수집 및 전처리

04

모델
학습 및 평가

05

관련 이슈
및
확장 방안

프로젝트 배경 및 필요성



실제 집의 책장 사진

1

서점에서 책 정보를 편리하게 검색할 수 있었으면 하는 바람

2

책장 안에 어떤 책이 있는지에 대한 호기심



여러 권의 책 제목 검출 및 웹 검색 서비스

프로젝트 배경 및 필요성



간편하게 책 정보 확인 가능

출판사, 작가, 다른 독자의 리뷰 및 평점 등의 정보를
기존의 검색 방식보다 편리하게 검색 가능



책 훼손 방지

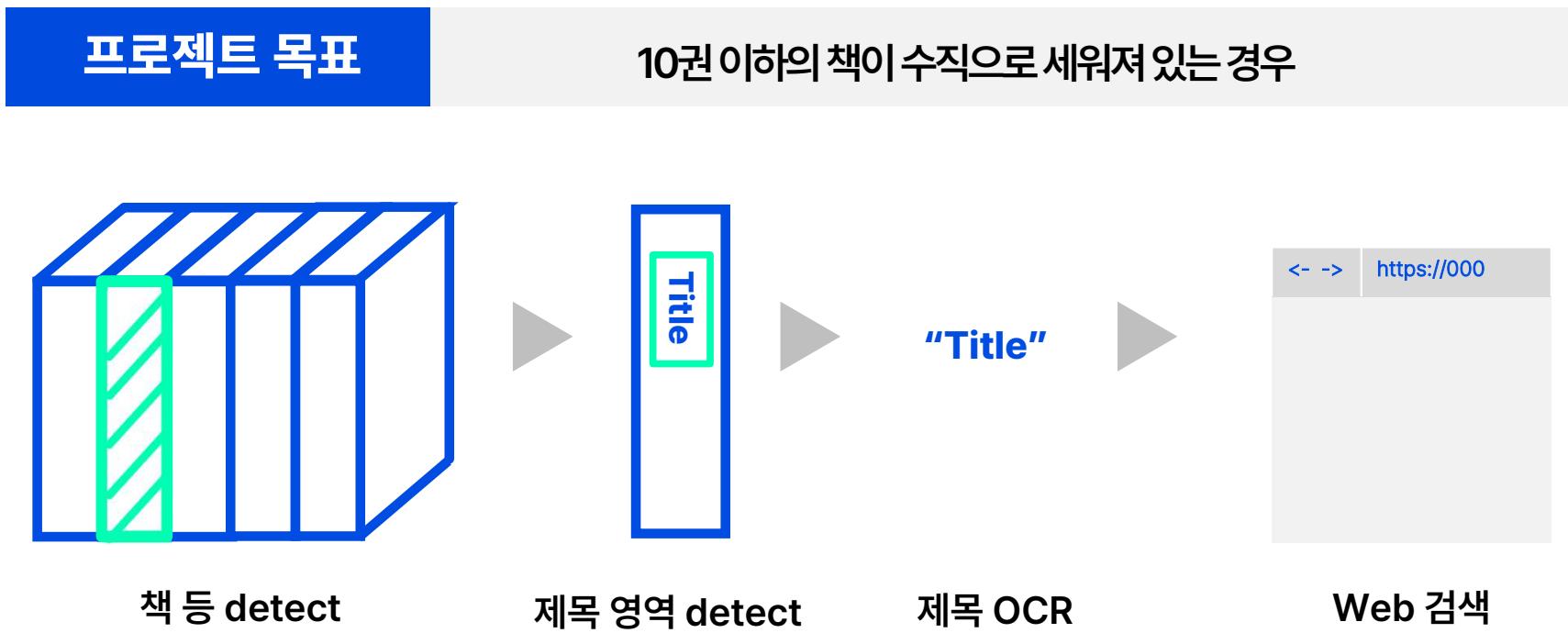
서점 및 도서관에서 책 선택 과정에서
도서 훼손 문제 발생



오프라인 매장에서 구매 시 활용

현재 많은 소비자가 활용하는 서비스
(교보문고 바로드림)

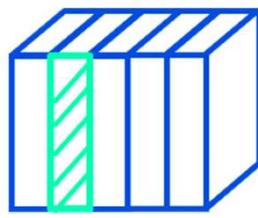
프로젝트 개요 및 목표



프로젝트 개요 및 목표



Stage 1



책등 detect



Stage 2



제목 영역 detect



Stage 3

"Title"

제목 OCR

"SF의 힘"

파이프라인

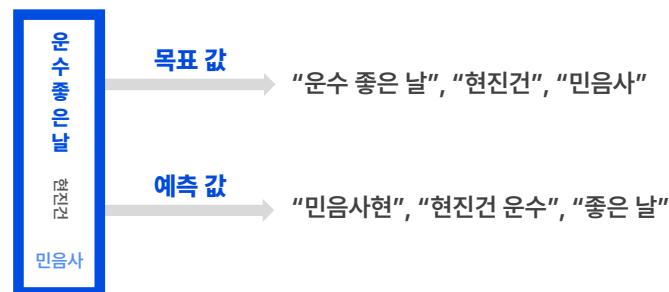
프로젝트 개요 및 목표

Two Stage Detection

사람이 책을 인식하는 방식과 유사한 과정으로 진행 [책장 → 책 등 → 책 제목]

Single Stage Detection 시 문제점

- Single Stage : 책등 detection
- OCR 실행 시 문자 인식 순서 문제로 인해 불규칙한 텍스트 검출
→ 책제목, 부제, 저자, 출판사 등이 혼재된 결과물



Two Stage Detection 시 장점

- Two Stage : 책등 Detection + 책제목 Detection
- 책제목 이외에 저자, 출판사 등의 정보를 추출할 경우에도 분리하여 그룹화하기에 용이
- 검출할 object를 줄임으로써 더 단순한 모델을 사용할 수 있음

프로젝트 개요 및 목표

핵심 기술

Object Detection OCR Web Application

활용 기술

YOLO, Google Cloud Vision, Streamlit, Pytorch, OpenCV, Pandas, Wandb

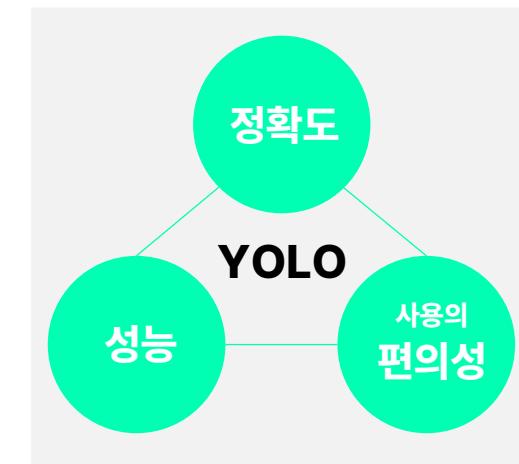
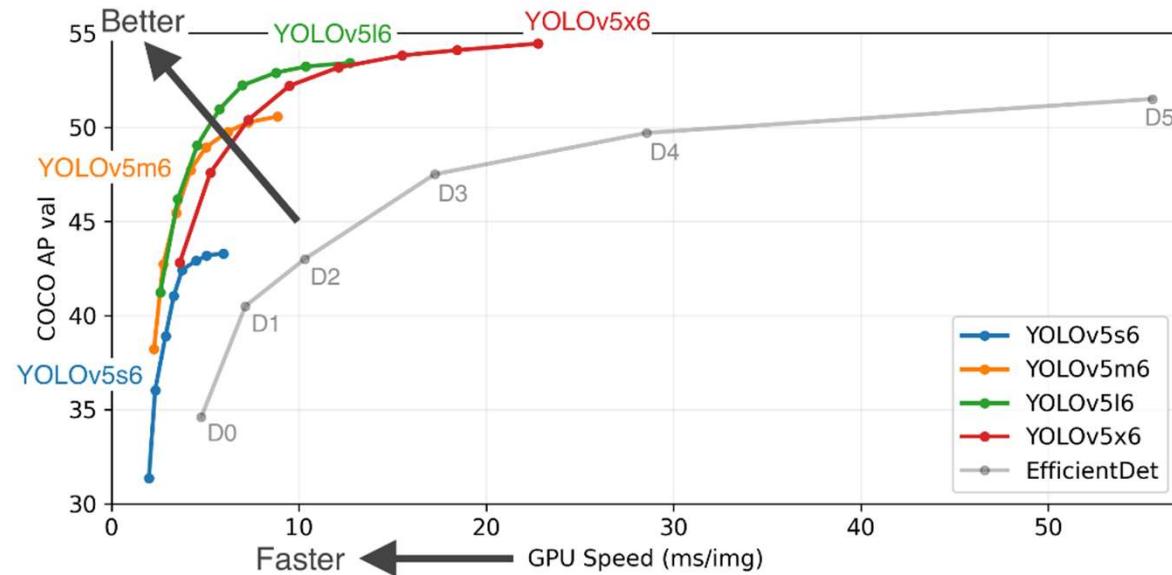
개발 환경

Colab Pro

프로젝트 개요 및 목표

WHY YOLO?

모바일 앱 구현 고려 → 모델이 가벼운 동시에 정확도가 높은 특징 필요



데이터 수집 및 전처리

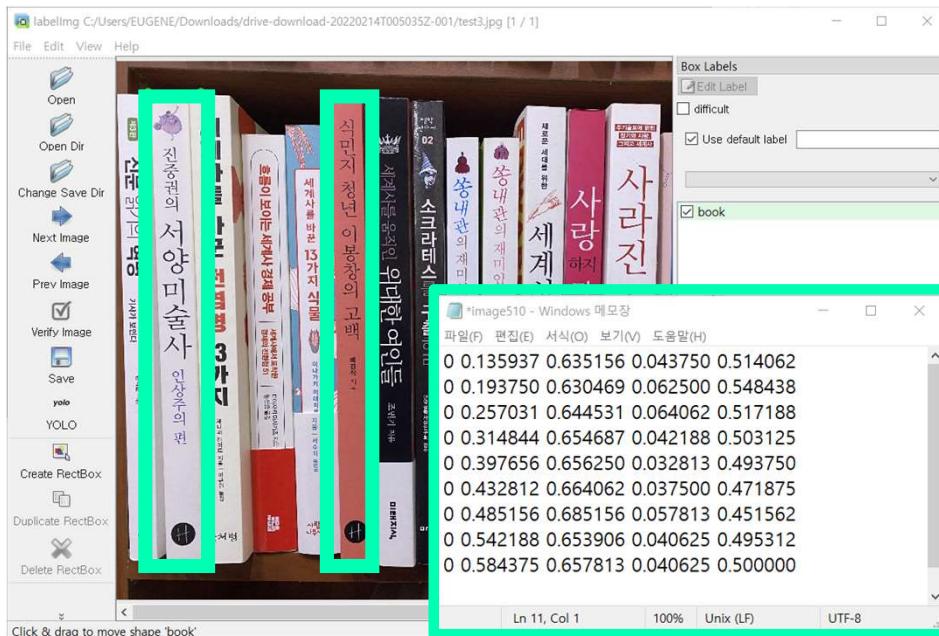


데이터 수집 및 전처리

데이터 전처리

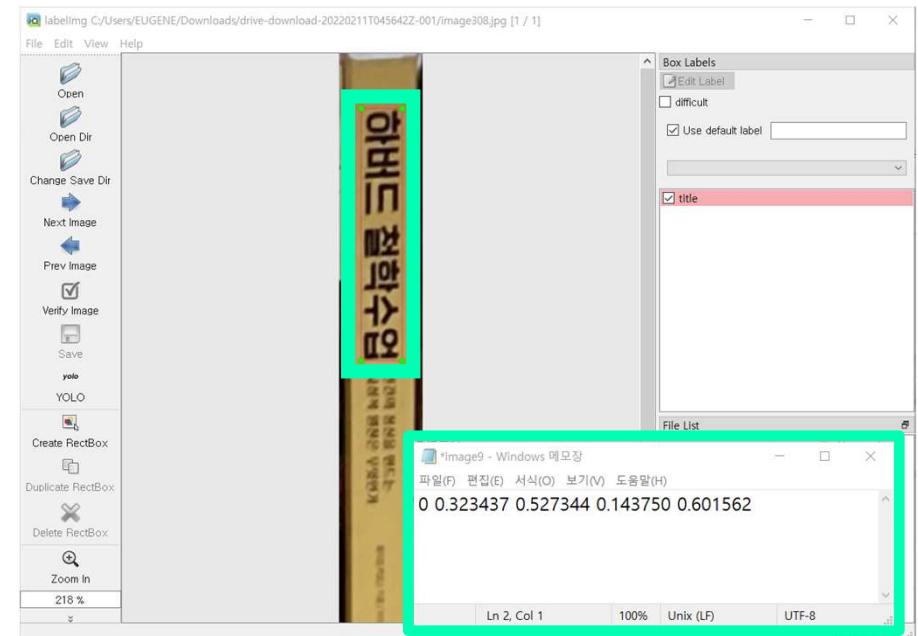
Labeling (LabelImg 사용)

책등 labeling



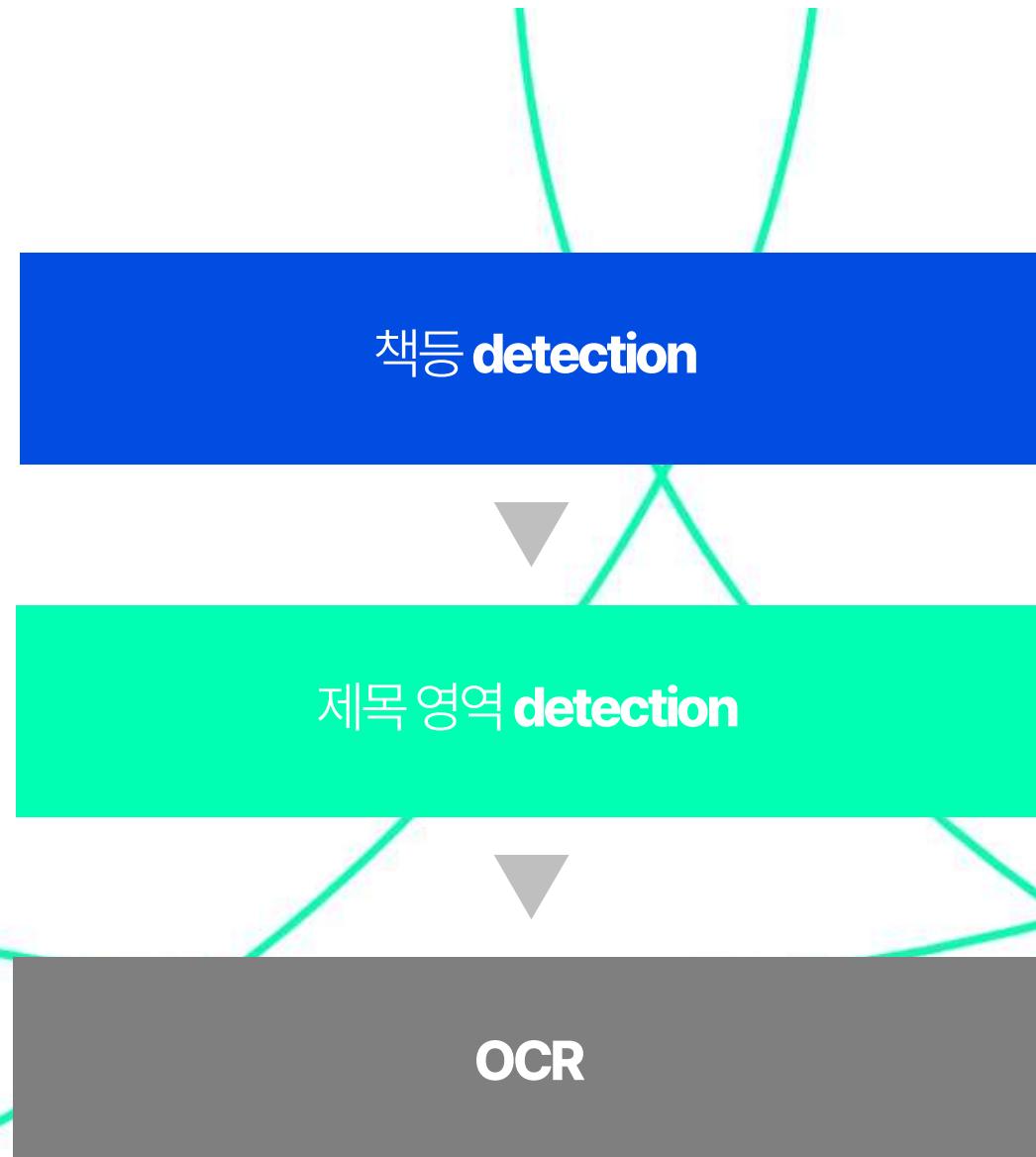
↑ Labeling 결과

책제목 labeling



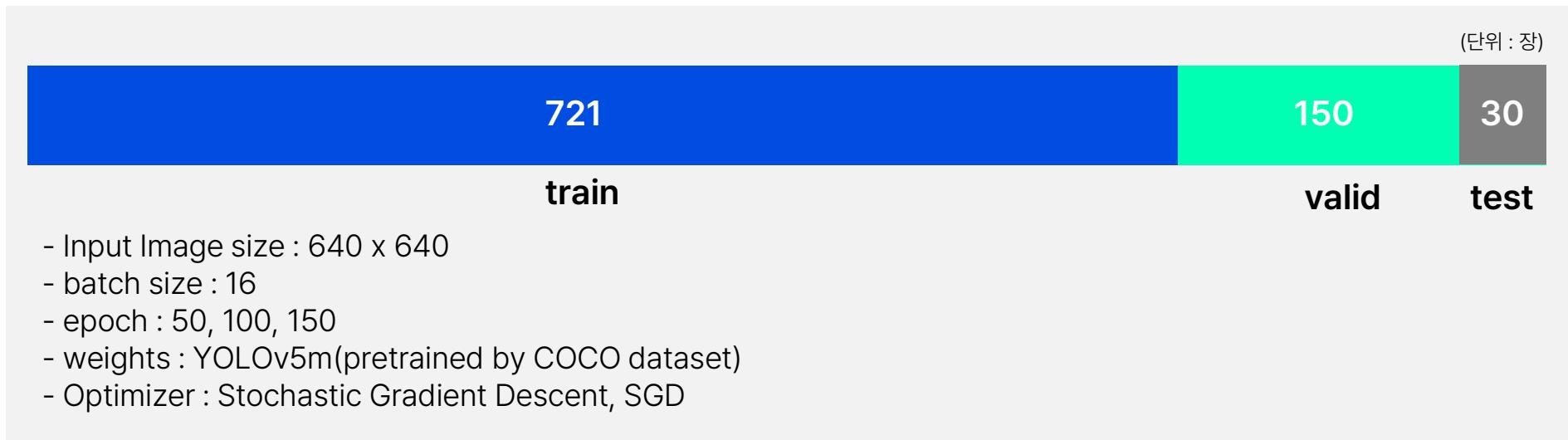
↑ Labeling 결과

Part 4 모델링 및 평가



모델링 및 평가 - (1) 책 등 detection

모델 Train

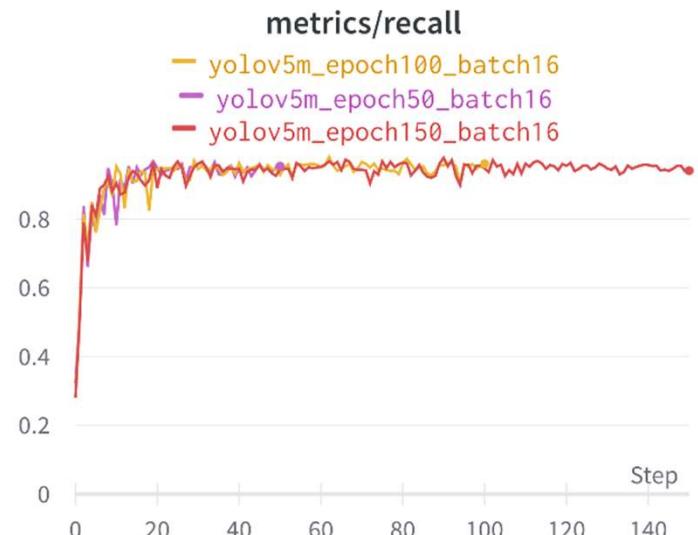
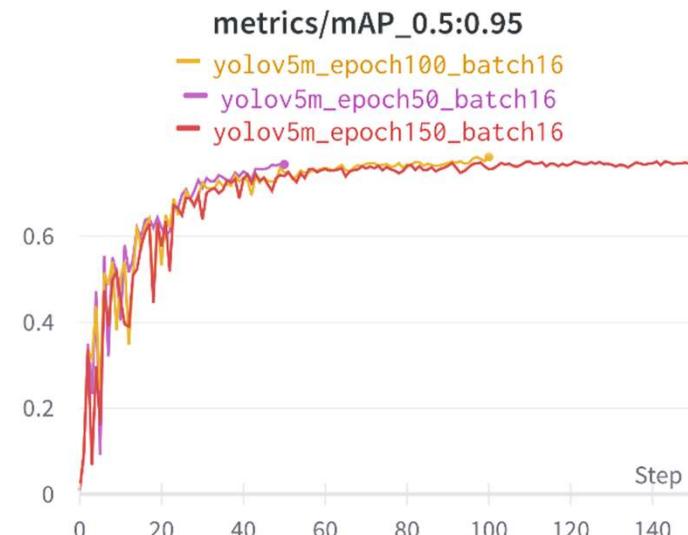
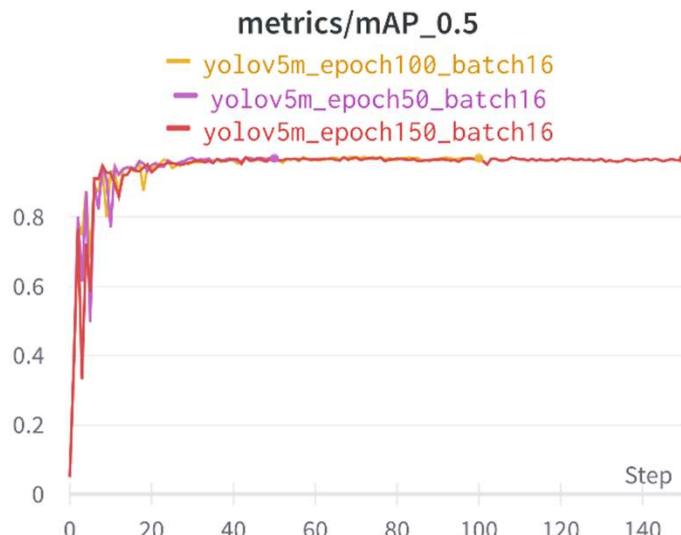


Epochs	Box loss	Obj loss	Cls loss	Precision	Recall	mAP (0.5)	mAP (0.5 ~ 0.95)
50	0.0281	0.06253	0	0.9275	0.9546	0.9699	0.7674
100	0.02568	0.05613	0	0.9254	0.9612	0.9704	0.7849
150	0.02454	0.05203	0	0.9385	0.9413	0.96354	0.774

모델링 및 평가 - (1) 책 등 detection

Ablation Study(1)

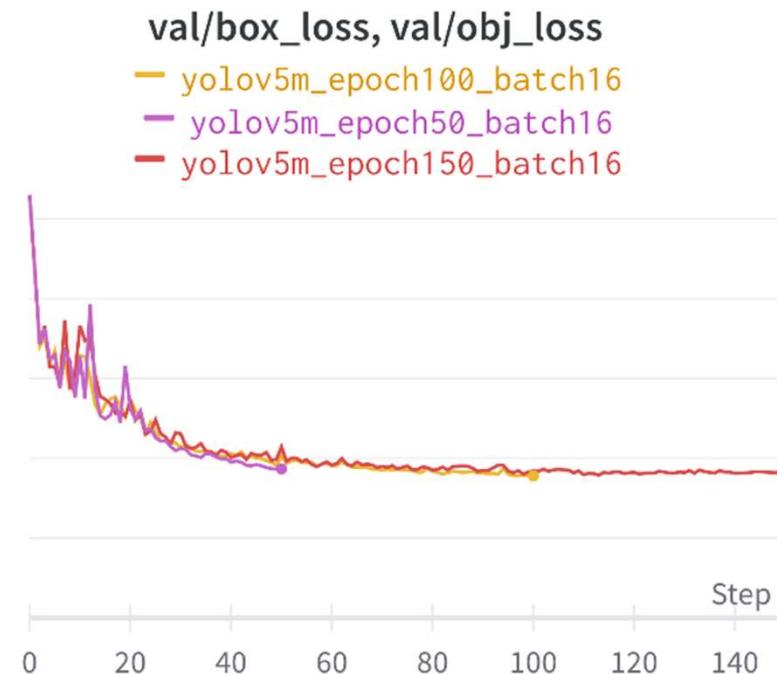
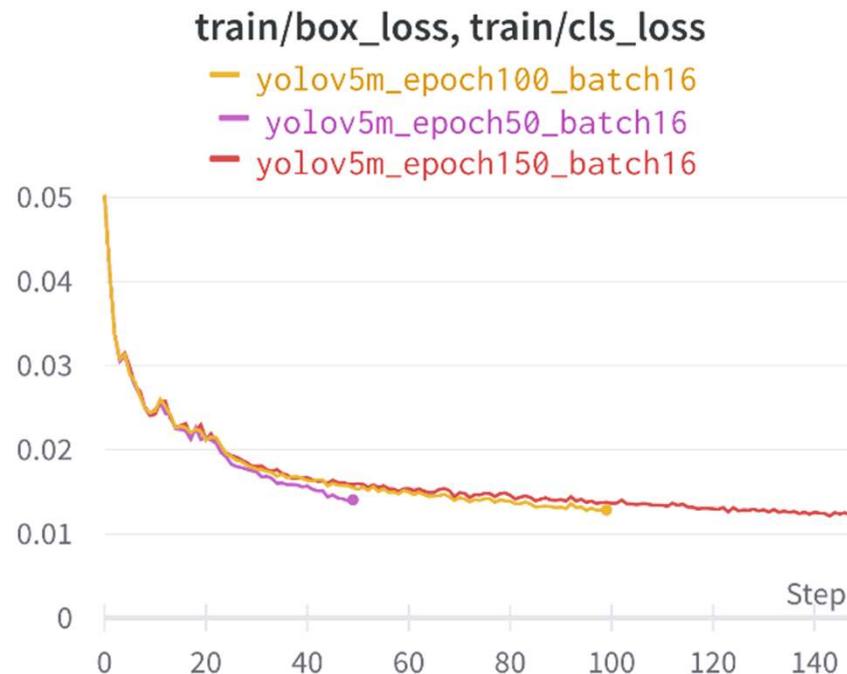
Batch size는 16로 동일, Epoch 50 vs 100 vs 150



모델링 및 평가 - (1) 책 등 detection

Ablation Study(1)

Batch size는 16로 동일, Epoch 50 vs 100 vs 150



모델링 및 평가 - (1) 책 등 detection

정량평가

목표 인식률 90%

Test dataset : 30장(517권)

$$\text{인식률(%)} = \frac{\text{검출한 bbox 개수}^*}{\text{실제 labeling된 bbox 개수 (실제 책 권수)}} \times 100$$

Confidence Threshold	Epochs 50	Epochs 100	Epochs 150	(인식률:%)
0.25	95.36%	91.49%	91.30%	
0.3	94.78%	90.91%	90.52%	
0.35	93.62%	89.94%	89.56%	
0.4	90.91%	89.17%	82.98%	

* 검출한 bbox 개수 = min(실제 책 권수, 검출한 bbox 개수)

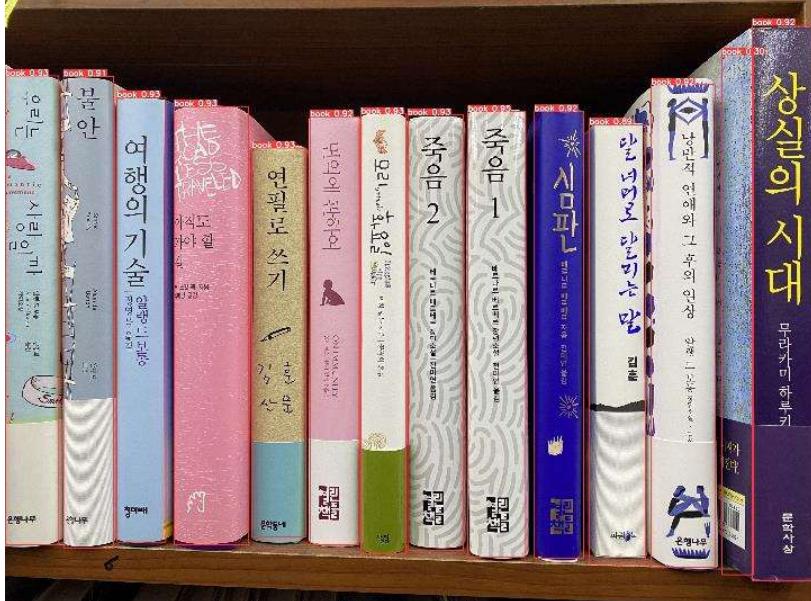
모델링 및 평가 - (1) 책 등 detection

정량평가

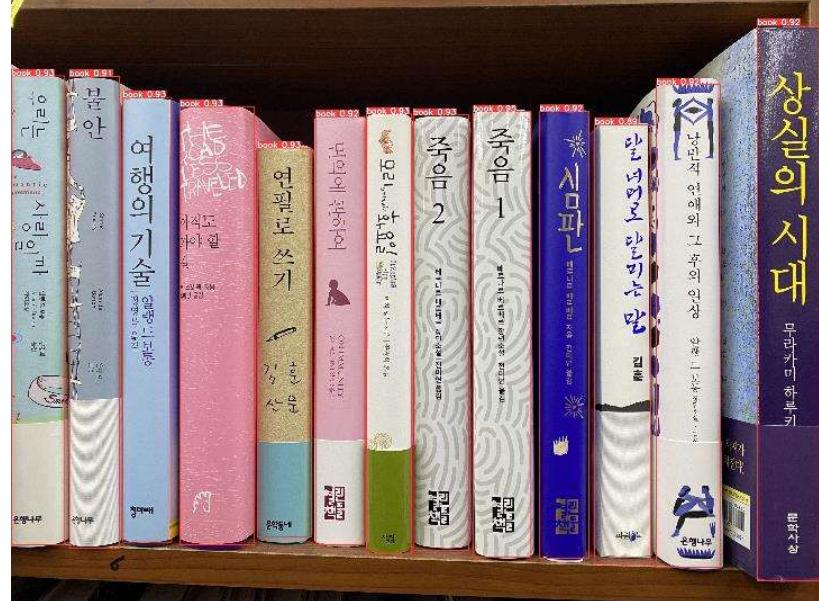
목표 인식률 90%

Test dataset : 30장(517권)

↓ Epochs 100, Confidence Threshold 0.3



↓ Epochs 100, Confidence Threshold 0.4



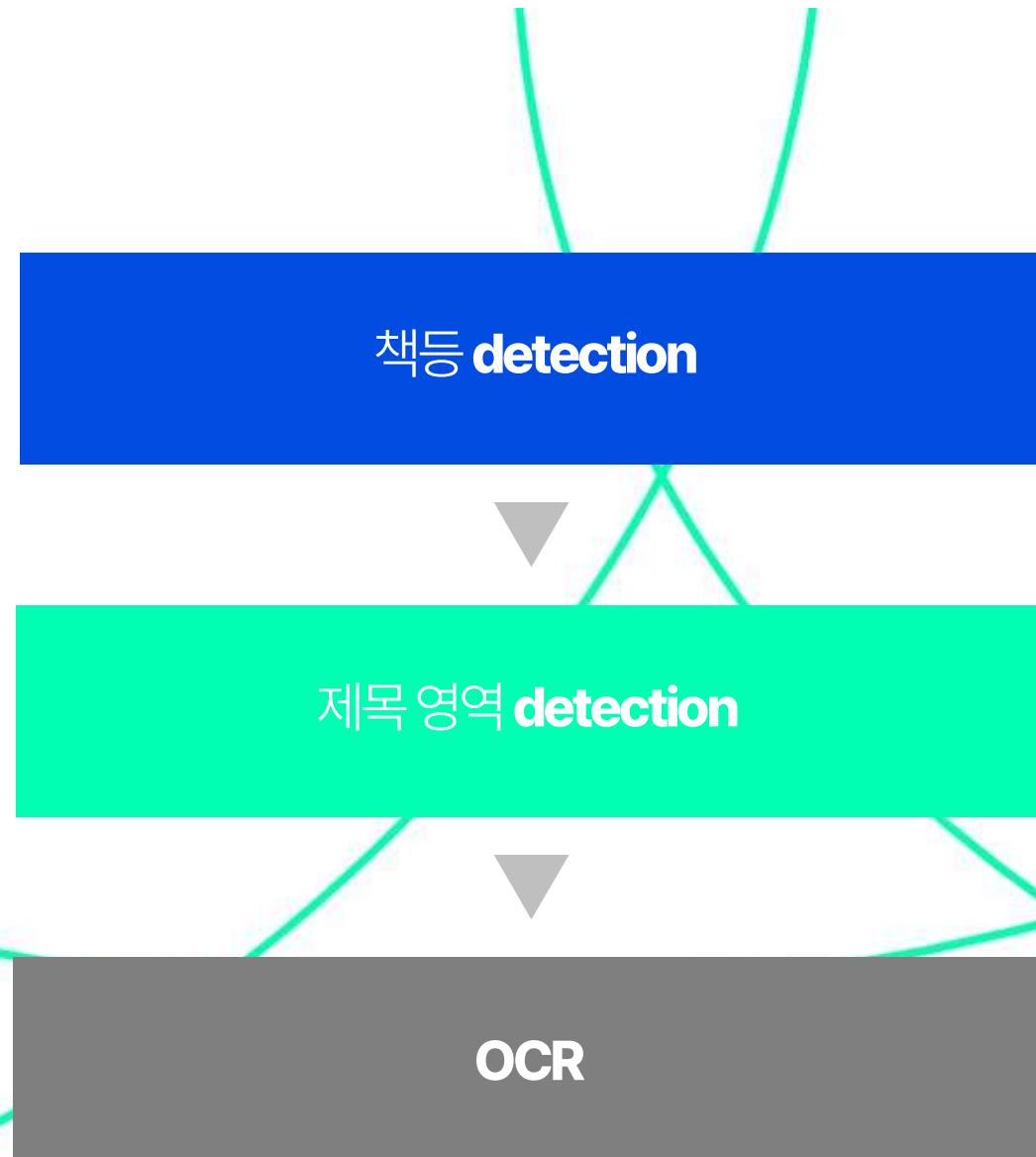
모델링 및 평가 - (1) 책 등 detection

정량평가

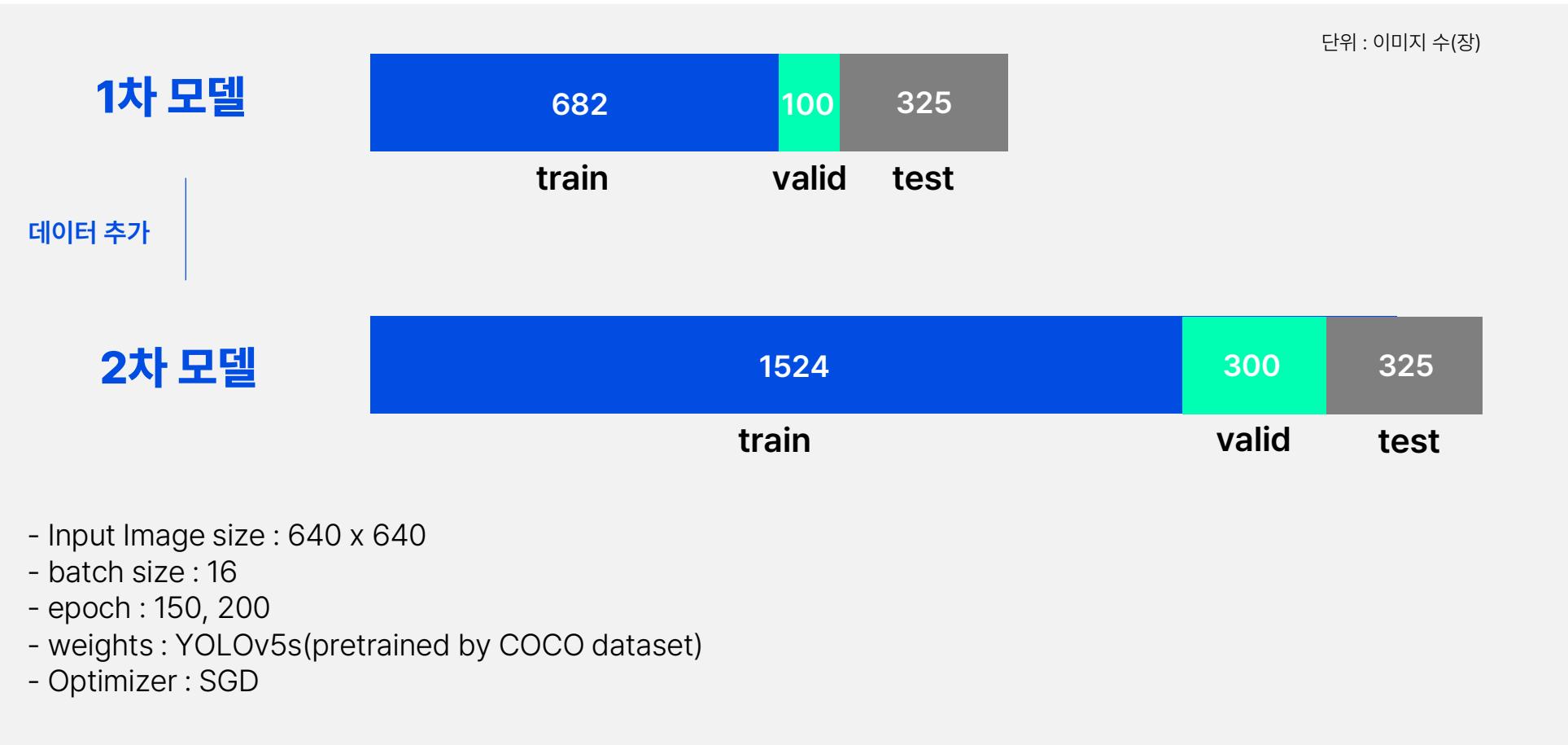
목표 인식률 90%

Epochs	Confidence threshold	총 오검출 개수	책 옆면을 책으로 인식	2권의 책을 1권으로 인식	1권의 책을 2권으로 인식
50	0.3	13	4	6	3
	0.35	10	2	6	2
	0.4	6	2	3	1
100	0.3	4	1	2	1
	0.35	2	0	1	1
	0.4	2	0	1	1
150	0.3	5	0	5	0
	0.35	4	0	4	0
	0.4	4	0	4	0

Part 4 모델링 및 평가

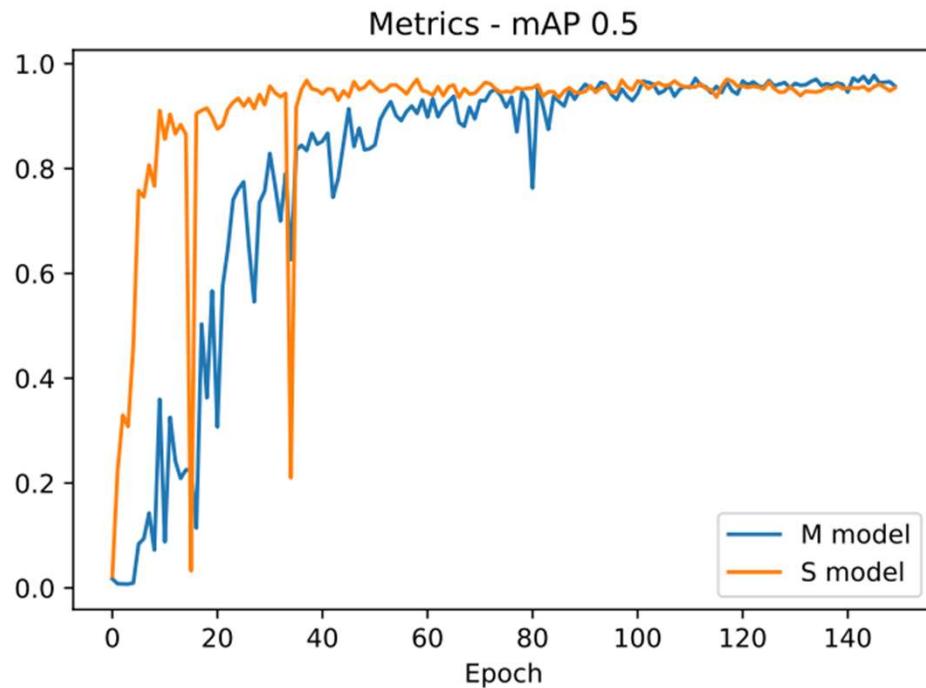


모델링 및 평가 - (2) 제목 detection

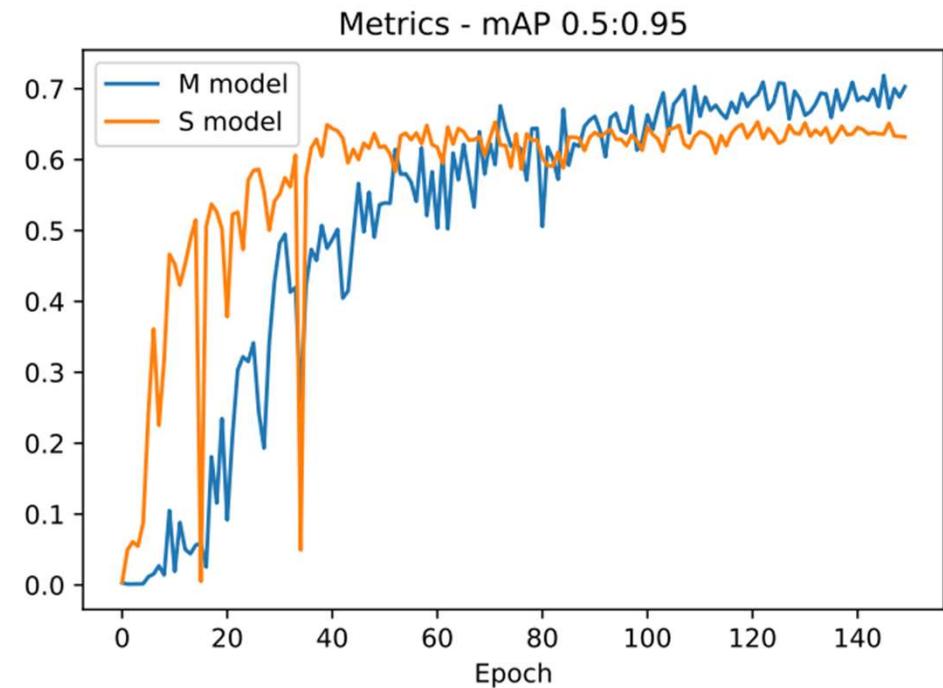


모델링 및 평가 - (2) 제목 detection

Ablation Study(2)



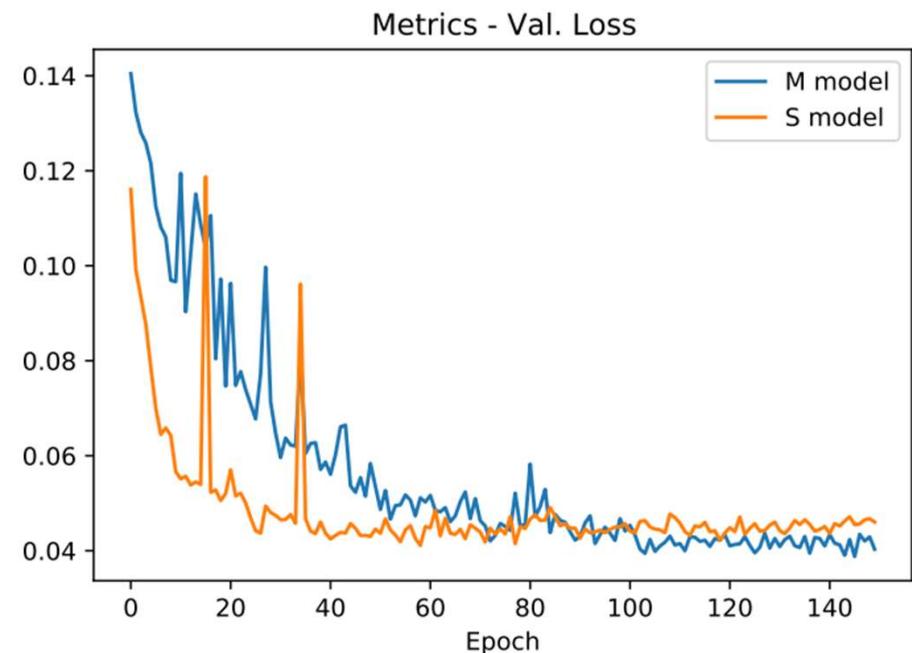
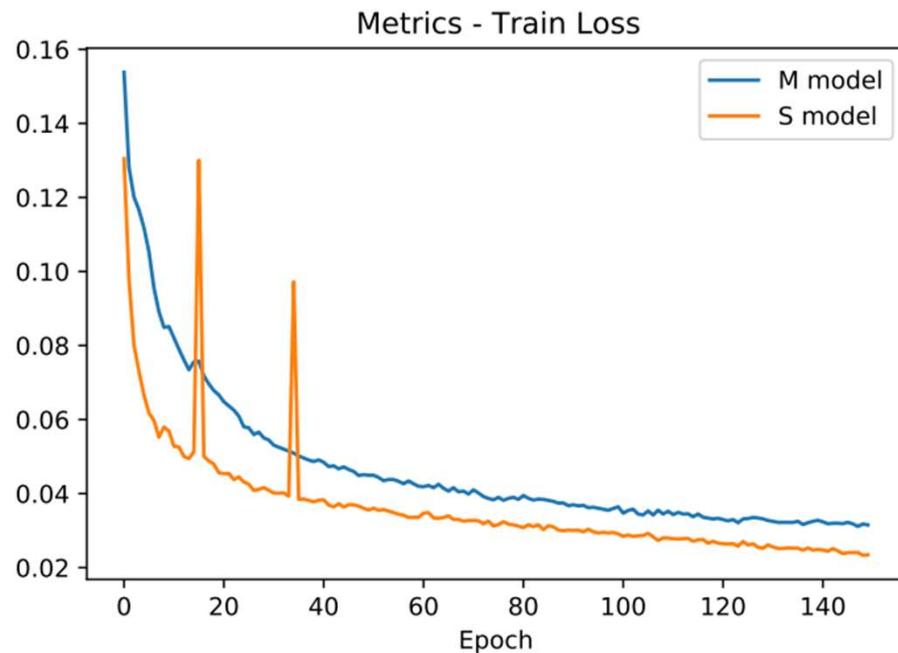
M 모델 vs S 모델



모델링 및 평가 - (2) 제목 detection

Ablation Study(2)

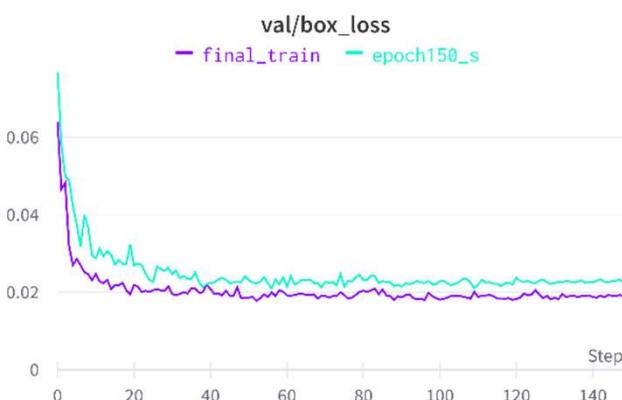
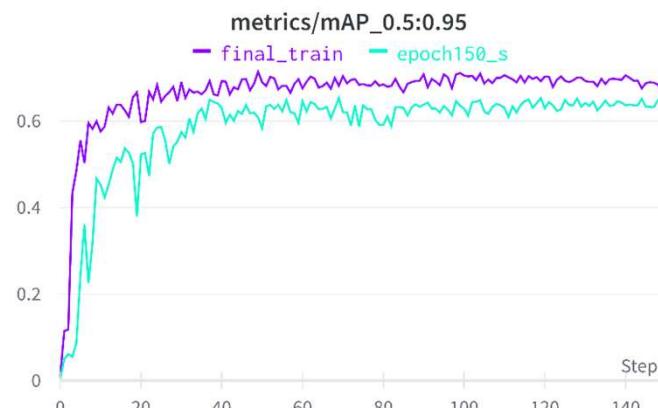
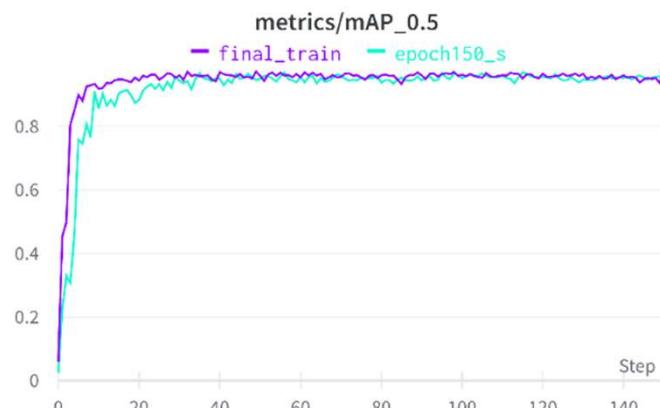
M 모델 vs S 모델



모델링 및 평가 - (2) 제목 detection

Ablation Study(2)

데이터 수에 따른 학습 과정 비교 : 682 VS 1524



모델링 및 평가 - (2) 제목 detection

정량 평가

목표 인식률 90%

Labeling한 bbox와 예측 bbox의 IOU 비교

- 테스트 데이터 : 325개
- S1 : 학습데이터 682개로 학습된 s모델
- S2 : 학습데이터 1524개로 학습된 s모델

$$\text{Intersection over Union} = \frac{\text{Intersection of Boxes}}{\text{Union of Boxes}}$$



모델링 및 평가 - (2) 제목 detection

정량 평가

목표 인식률 90%

IoU 범위에 따른 bbox 검출 개수		
IoU	S1.	S2.
0.9 – 1.0	154	175
0.8 – 1.0	244	256
0.7 – 1.0	265	268
0.6 – 1.0	275	281
0.6 미만	37	41
total	312	322

IoU 범위에 따른 검출 비율		
IoU	S1.	S2.
0.9 – 1.0	47.38%	53.85%
0.8 – 1.0	75.08%	78.77%
0.7 – 1.0	81.54%	82.46%
0.6 ~ 1.0	84.62%	86.46%

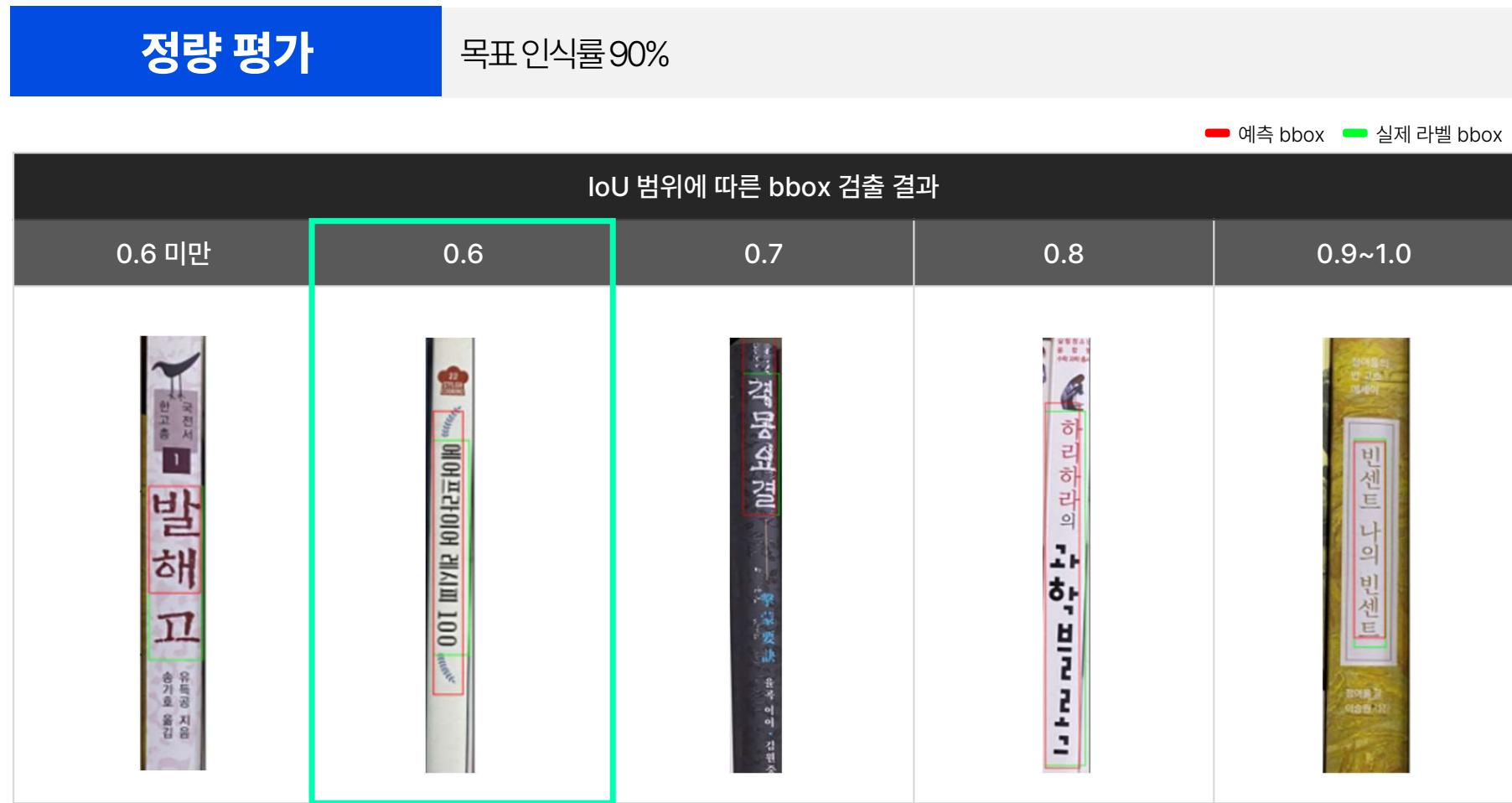
모델링 및 평가 - (2) 제목 detection

정량 평가

목표 인식률 90%

■ 예측 bbox ■ 실제 라벨 bbox

IoU 범위에 따른 bbox 검출 결과



모델링 및 평가 - (2) 제목 detection

정량 평가

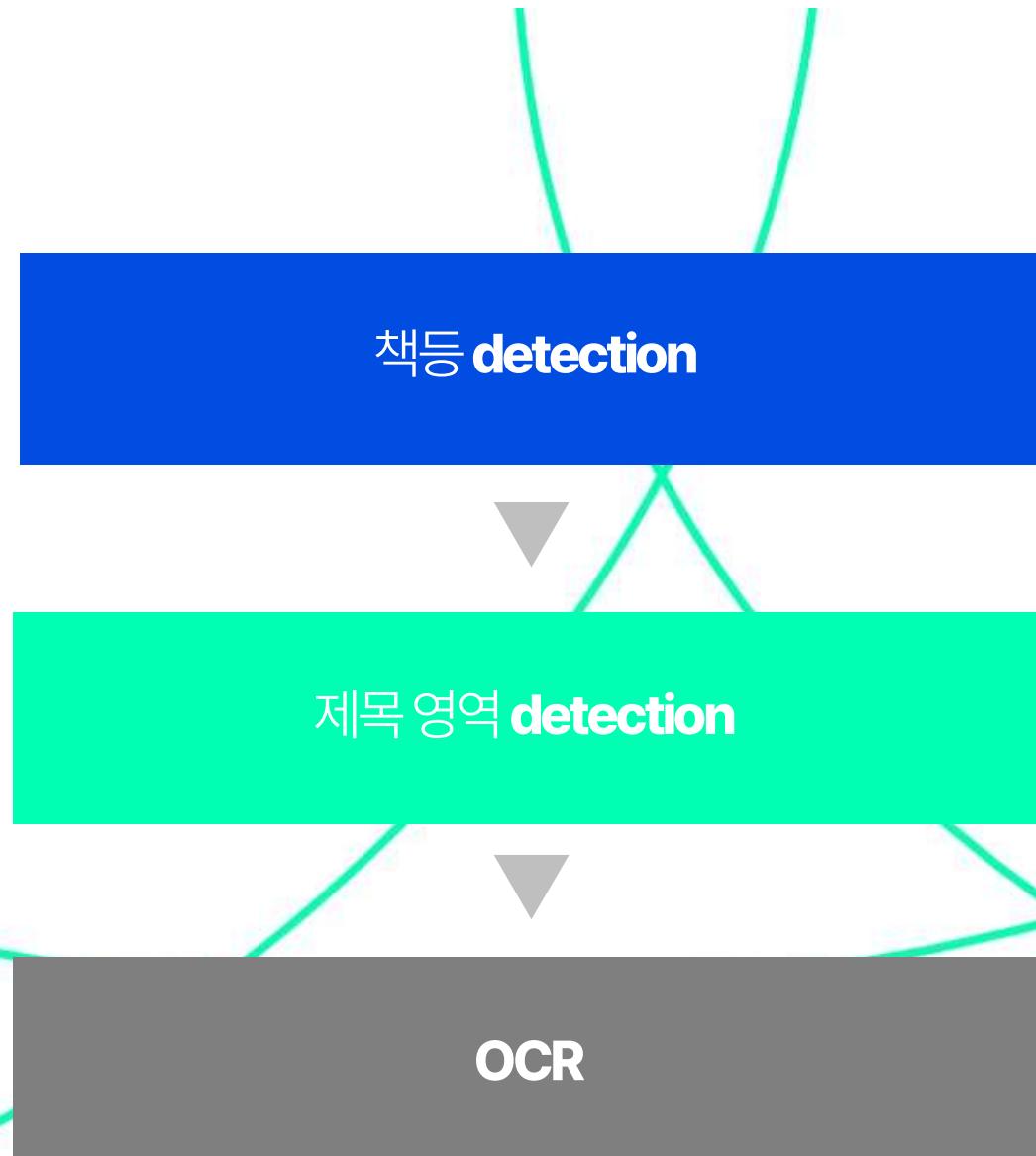
목표 인식률 90%

IoU 범위에 따른 bbox 검출 개수		
IoU	S1.	S2.
0.9 – 1.0	154	175
0.8 – 1.0	244	256
0.7 – 1.0	265	268
0.6 – 1.0	275	281
0.6 미만	37	41
total	312	322

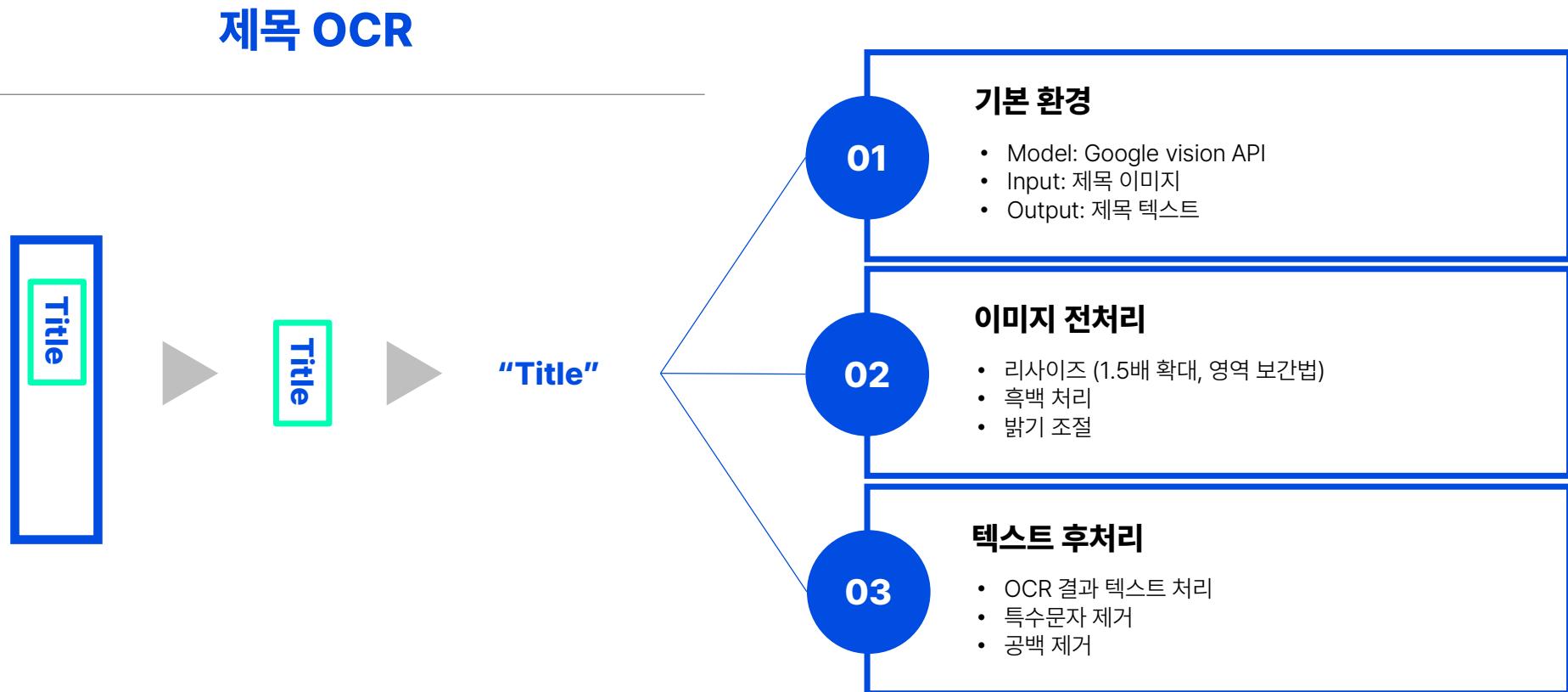
IoU 범위에 따른 검출 비율		
IoU	S1.	S2.
0.9 – 1.0	47.38%	53.85%
0.8 – 1.0	75.08%	78.77%
0.7 – 1.0	81.54%	82.46%
0.6 ~ 1.0	84.62%	86.46%

Part 4

모델링 및 평가



모델링 및 평가 - (3) 제목 OCR



모델링 및 평가 - (3) 제목 OCR

OCR 모델 비교

- 동일한 제목 이미지 10장으로 테스트 진행
- Naver API, Google API, Kakao 3사 OCR 서비스 비교
- 인식 결과가 50% 이상인 Naver API와 Google vision API 상세 비교

Naver API		Google vision API
결과	50%	80%
장점	<ul style="list-style-type: none">자연스러운 단어 구분다양한 폰트 인식	<ul style="list-style-type: none">가로-세로 인식 안정적대부분의 글자 인식
단점	<ul style="list-style-type: none">가로 회전된 글자 인식 불가	<ul style="list-style-type: none">디자인 폰트, 필기체의 경우 인식률 감소

모델링 및 평가 - (3) 제목 OCR

정상 검출

모델	결과
Naver	냉정과 열정사이
Google	냉정 과 열정 사이
kakao	



Naver



Google

디자인 폰트 인식

모델	결과
Naver	운수 좋은 날
Google	S P A 노 은 은 날
kakao	날



Naver



Google



Kakao

제목이 회전된 경우

모델	결과
Naver	인한
Google	단 테 의 신 곡 살 인
kakao	

글자가 회전



Naver



Google

모델링 및 평가 - (3) 제목 OCR

정량 평가

목표 정확도 80%

기본 환경

- **테스트 데이터**: 제목 크롭 이미지 380장
- **모델**: Google Vision API
- **적용한 이미지 전처리**
 - ① 리사이즈 (1.5배, 영역 보간, 바이크ubic 보간)
 - ② 흑백 변환
 - ③ 대비 조절
 - ④ 밝기 조절

평가 방법

- **Step1.** difflib 라이브러리의 SequenceMatcher 사용하여 실제 제목과 예측 값을 비교하여 정확도 계산 (0~100%)
- **Step2.** 오차 범위별 정확도 계산

$$\frac{\text{기준오차범위내 데이터수}}{\text{전체데이터수}} \times 100$$

오차 범위	상세 기준
0 %	정확도 100% 데이터
0 ~ 10 %	정확도 90% 이상 데이터
0 ~ 20 %	정확도 80% 이상 데이터

모델링 및 평가 - (3) 제목 OCR

정량 평가

목표 정확도 80%

Step1. 실제 제목과 예측 값 비교

TARGET	운수좋은날
PREDICT	훈수좋은날
F T T T T	

80%

Step2. 오차 범위별 정확도 계산

구분	데이터 수	구분	정확도
정확도 100%	282	오차 범위 0%	74%
정확도 90% 이상	294	오차 범위 0~10%	85%
정확도 80% 이상	324	오차 범위 0~20%	88%
전체 데이터	380		

평가 방법

- Step1. difflib 라이브러리의 SequenceMatcher 사용하여 실제 제목과 예측 값을 비교하여 정확도 계산 (0~100%)
- Step2. 오차 범위별 정확도 계산

$$\frac{\text{기준 오차 범위 내 데이터 수}}{\text{전체 데이터 수}} \times 100$$

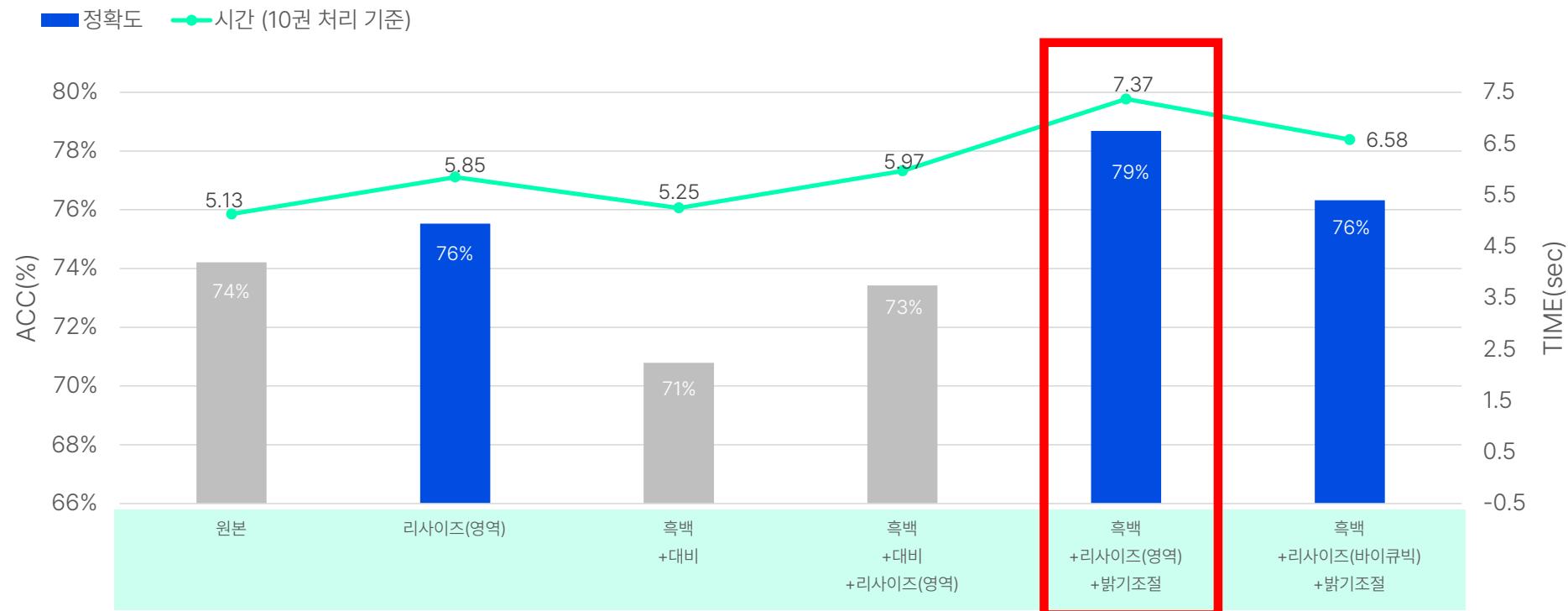
오차 범위	상세 기준
0 %	정확도 100% 데이터
0 ~ 10 %	정확도 90% 이상 데이터
0 ~ 20 %	정확도 80% 이상 데이터

모델링 및 평가 - (3) 제목 OCR

전처리 기법에 따른 성능 비교*

흑백처리, 리사이즈(1.5배, 영역 보간), 밝기조절을 적용 후 원본 대비 정확도 5%p 및 처리속도 2.24 sec 증가

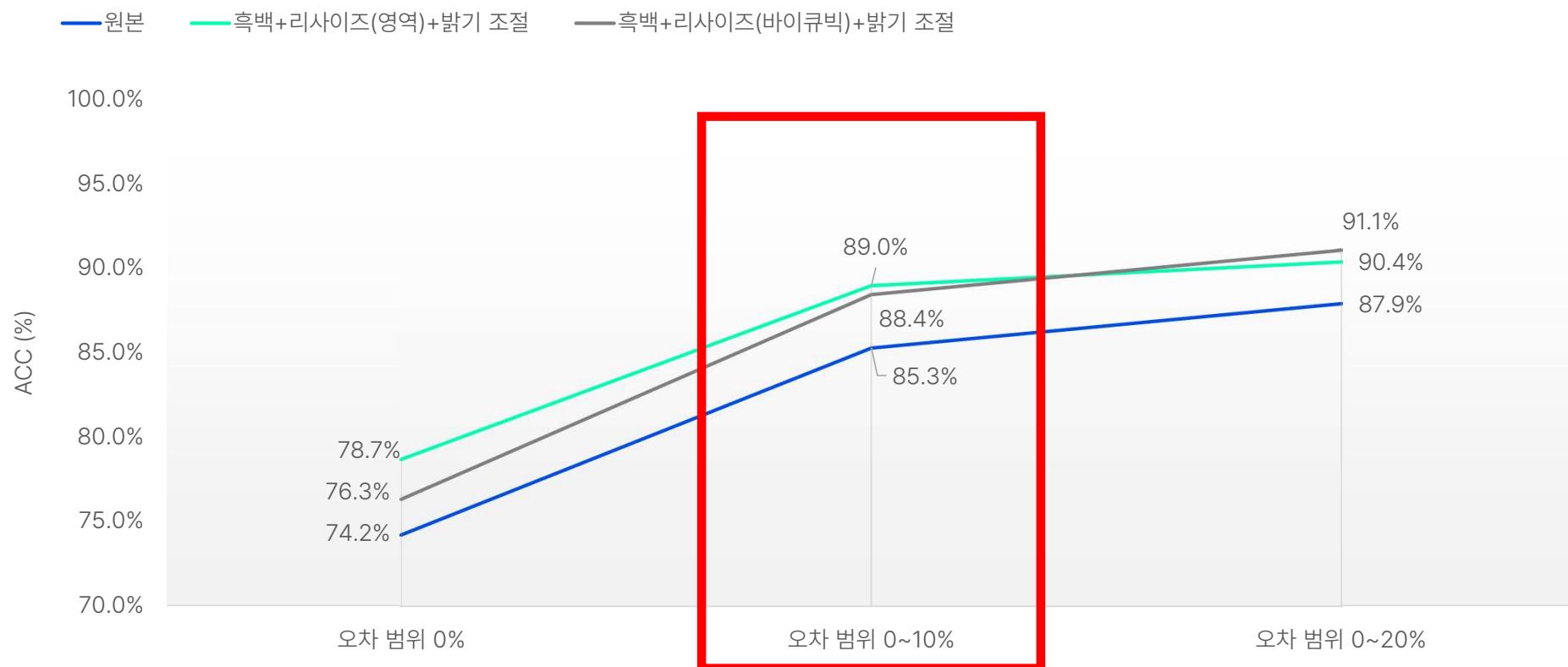
* 정확도는 오차범위 0% 데이터 기준, 처리속도는 책 10권 기준으로 비교



모델링 및 평가 - (3) 제목 OCR

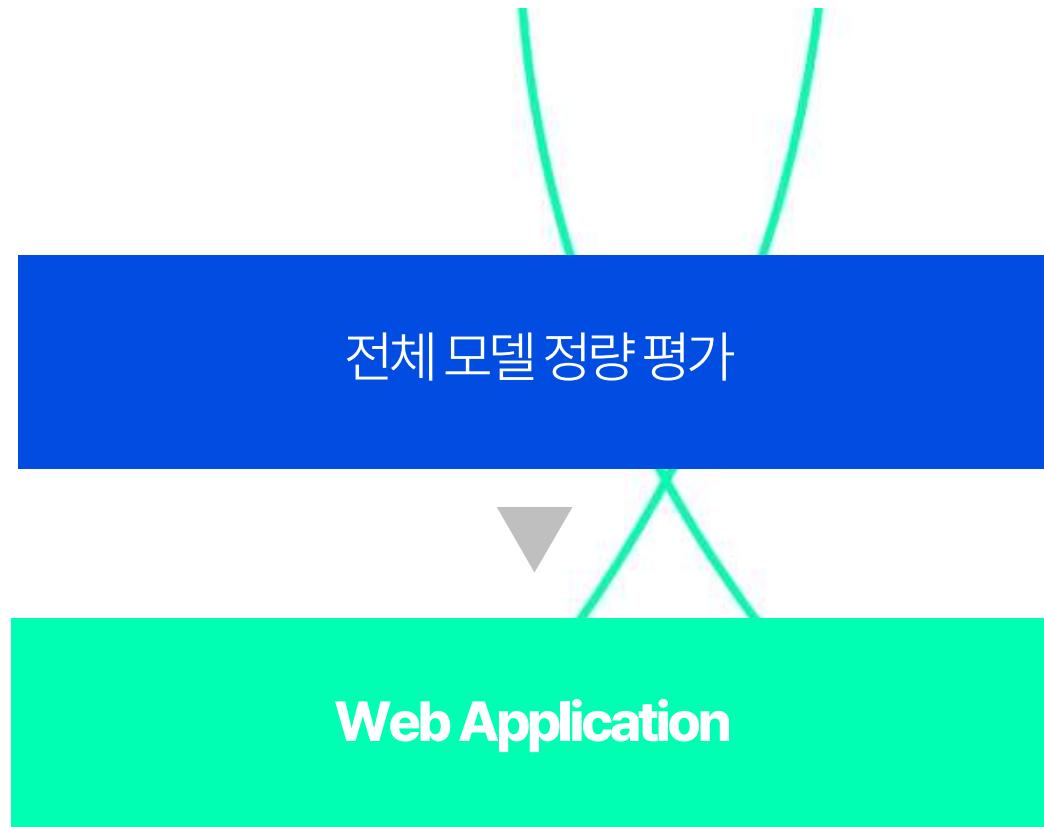
오차 범위에 따른 성능 비교

오차 범위 10%의 결과까지 고려할 경우 목표 정확도인 80% 이상인 약 85%, 89%, 88%의 정확도를 보임



Part 4

모델링 및 평가



모델링및평가- (4) 전체 모델 정량평가

Inference Time

Stage1에서 Stage3까지의 inference time 측정

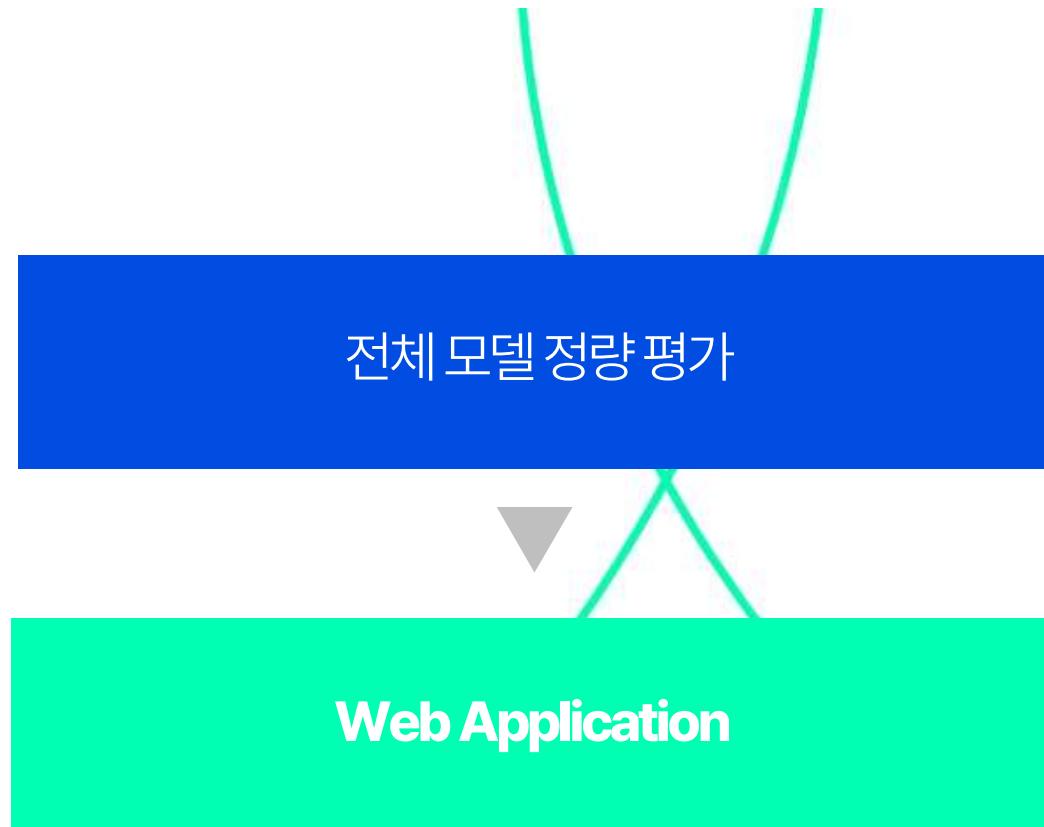
* stage3 시간은 이미지 전처리시간까지 포함된 결과

단위:초(sec)

	1권	2권	3권	4권	5권
Stage1	0.14	0.12	0.14	0.14	0.15
Stage2	0.01	0.02	0.02	0.03	0.03
Stage3	0.79	1.52	2.5	3.36	3.82
Total	0.95	1.67	2.67	3.54	4.01

	6권	7권	8권	9권	10권
Stage1	0.14	0.15	0.15	0.15	0.16
Stage2	0.04	0.04	0.05	0.06	0.06
Stage3	4.73	5.57	5.82	6.63	7.76
Total	4.92	5.78	6.04	6.85	8

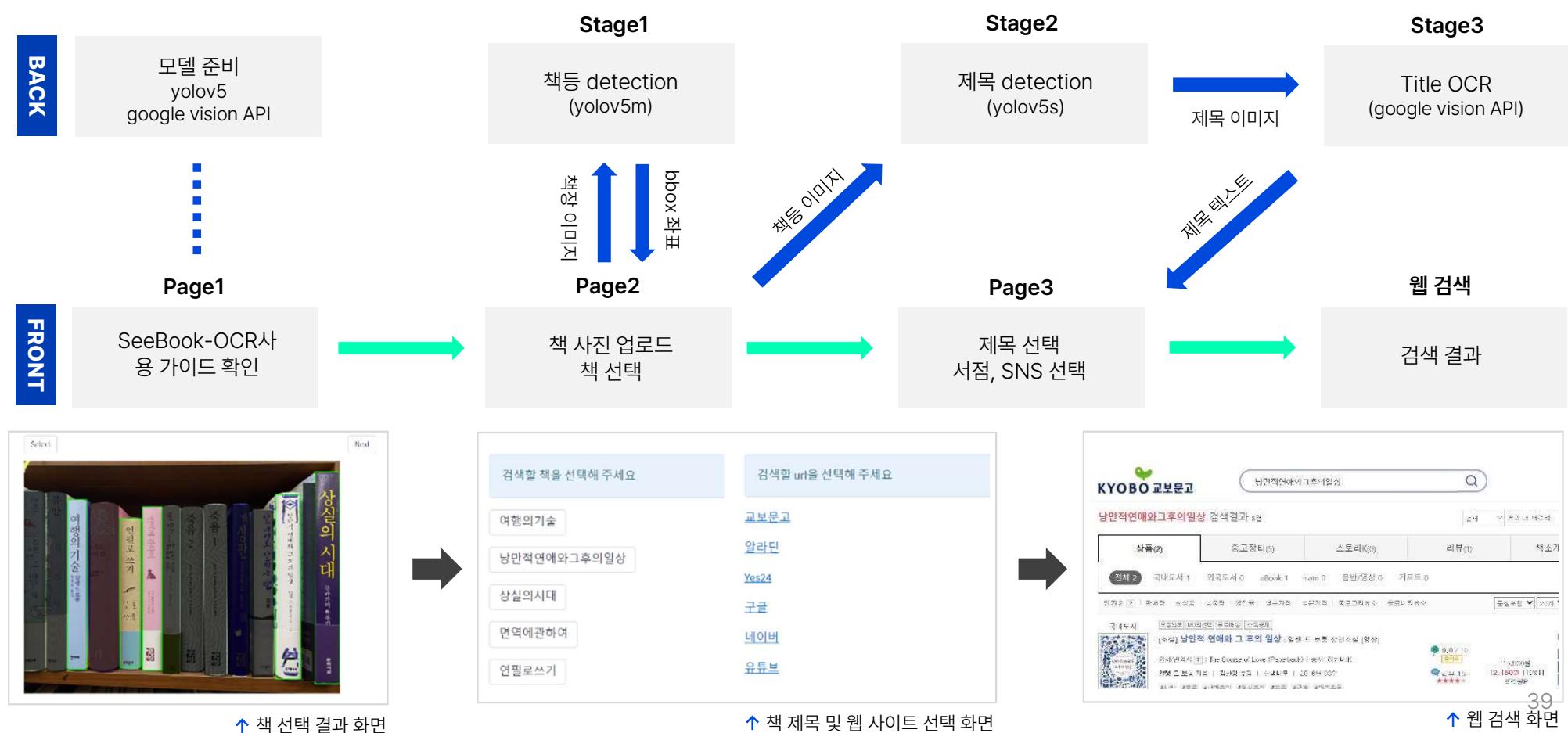
Part 4 모델링 및 평가



모델링 및 평가 - (5) Web Application

웹 애플리케이션

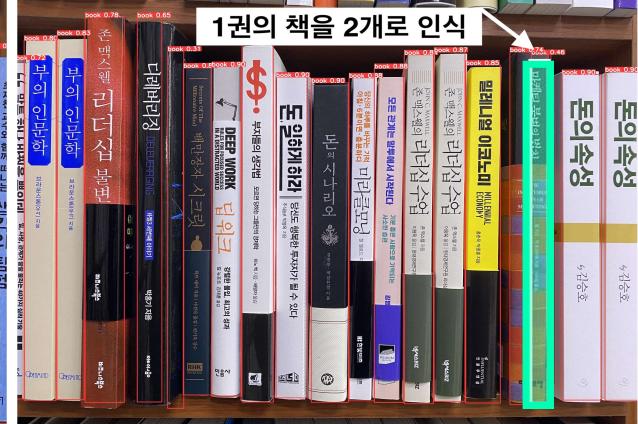
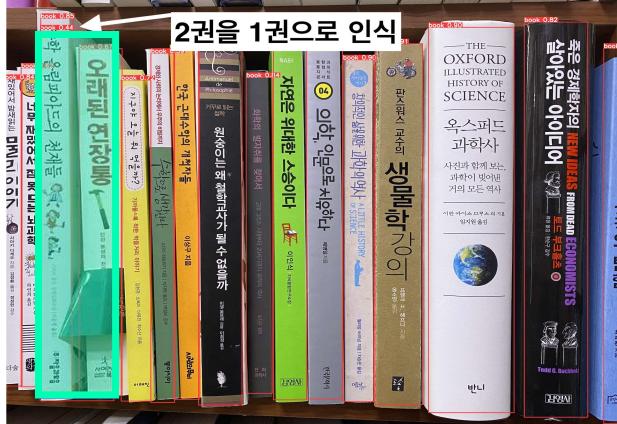
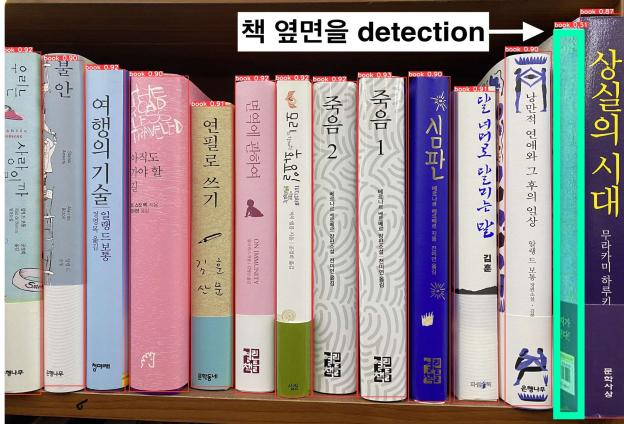
'Streamlit' 프레임워크를 사용하여 웹 애플리케이션을 제작 → 사용자의 편의를 위한 검색 서비스 제공



관련이슈

오검출 문제

책의 옆면 detection, 책 2권을 1권으로 인식, 1권의 책을 2권으로 detection하는 문제



해결방안

원하는 책만 선택하게 함으로써 오검출된 책은 무시

Conf. threshold 값을 조정하여 오검출된 bounding box는 제거

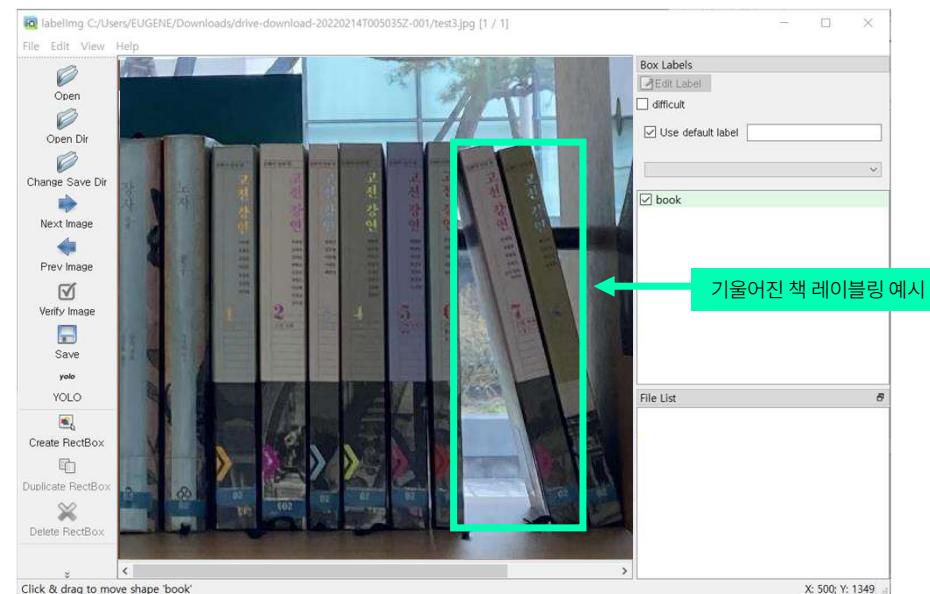
관련 이슈

기울어진 책은 탐지하지 못함

문제 원인

- Labelimg에서 수직 bounding box만을 지원
- 기울어진 책을 수직 bounding box로 labeling 시

여러 권의 책이 겹쳐지는 문제 발생



관련 이슈

기울어진 책은 탐지하지 못함

해결방안

이미지 촬영에 대한 [가이드라인](#)을 제공

[예시]

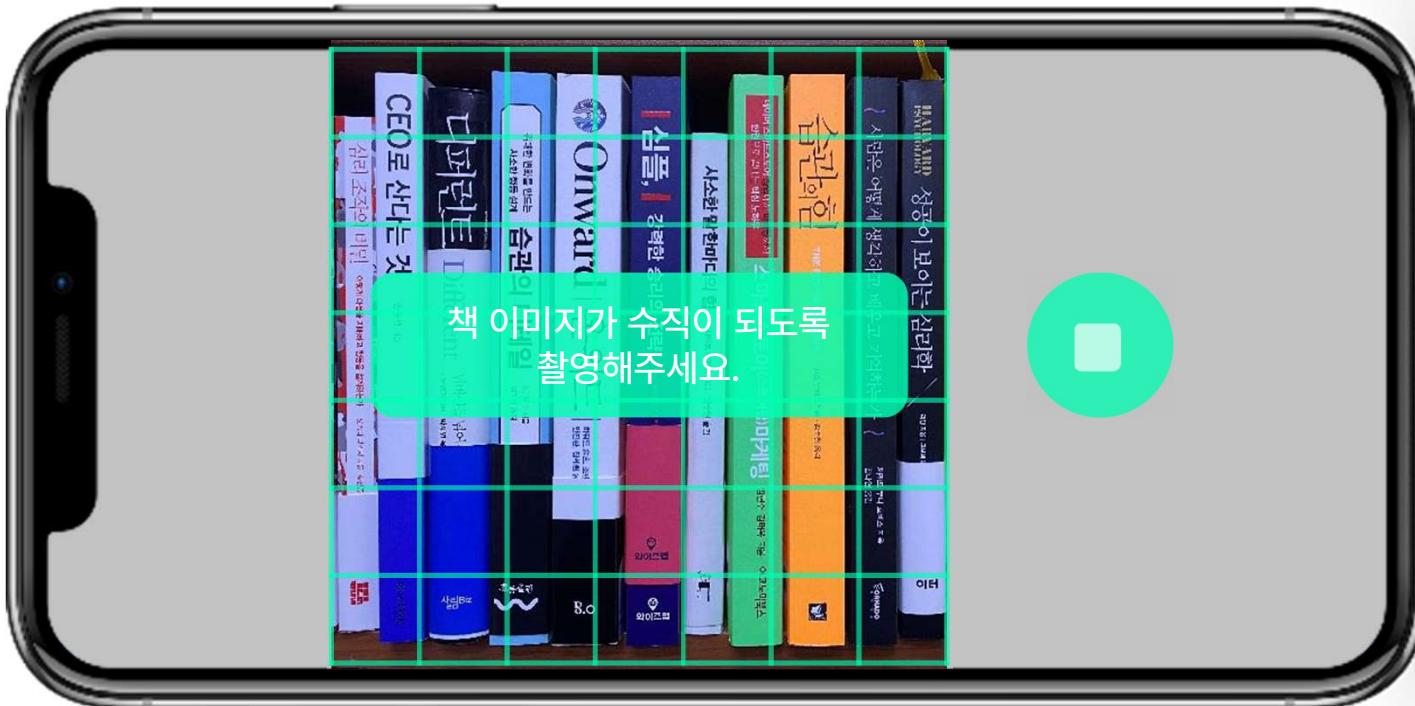
권장 사항 : 가로 촬영, 1단 책장, 수직 이미지, 10권 이내

주의 사항 : 2단 책장X, 기울어진 책X, 빛 반사 주의



관련 이슈

이미지 촬영에 대한 **가이드라인**을 제공

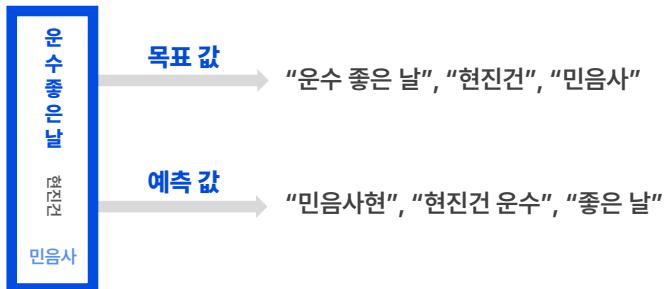


관련 이슈

OCR

문제 분석

- 책등 사진에서 OCR 진행 시 불규칙한 순서로 텍스트 검출
- OCR 결과 음절 일치율 저조
- 상대적으로 느린 처리 속도 문제 → 10권 기준 평균 6sec



책제목	result	compare1
일리아스	아	Fail
천국의열쇠	천국의열쇠	Fail
4대비극	가대비극FI	Fail
블랙스완		Fail
진실의흑역사	진실의흑역사K	Fail
세계사를바꾼10가지약	세계사를바꾼가지약	Fail
왜버스는한꺼번에오는걸까	TO오는걸	Fail
재미있는영재들의수학퍼즐2	영재들의수학퍼즐2재미있는	Fail

관련 이슈

OCR 관련 이슈

해결방안

- Stage2 단계를 통해 title detection 과정을 추가하여 불규칙적으로 텍스트를 검출하는 문제 해결
- 이미지 리사이즈, 흑백 및 밝기 조절 등의 전처리를 통한 OCR 성능 개선
- 사용자에게 원하는 책을 선택할 수 있게 하여 OCR 처리 속도 단축
 - 처리속도 개선 및 사용자의 니즈에 맞게 서비스 제공 가능

확장 방안

확장 방안

- **기울어진 책 detection**
 - Rotate-Yolov5
 - roLabelimg
- **저자, 출판사 등의 책 정보 추가 detection**
- **OCR 모델 제작**
 - 디자인 폰트에 대한 인식률 향상
 - 실행 시간 단축 (batch inference 사용)

localhost:8505

Select your page

Page1
 Page2
 Page3
 Page4

Page 1

SeeBook OCR

Object detection과 OCR을 활용한
여러권의 책 제목 검출 및 웹검색 서비스입니다.

책장 사진을 촬영하시면, 원하는 책에 대해서 책 제목을 텍스트로 바꿀 수 있으며
온라인 서점 및 SNS 등 책 정보 연결 페이지로 이동할 수 있습니다.

촬영 예시



좋은 촬영 예시



나쁜 촬영 예시

촬영 주의사항



수직 이미지



빛 반사 주의

10 ↓

10권 이내

1단

1단 책장

Done!

About 책봐조

팀장
박진우



총괄

- OCR 적용
- OCR 성능 평가
- Modeling(책등)
- Detection성능 평가
- Web app (page3)

팀원
김성연



OCR 팀

Main task

- OCR 적용
- OCR 성능 평가
- 이미지 전처리
- Web app(page2.2)

Support

- modeling(책등)

팀원
나웅



Detection 팀

Main task

- modeling(제목)
- Detection성능 평가
- Web app(통합)

Support

- 이미지 전처리
- modeling(책등)

팀원
장유진



Detection 팀

Main task

- modeling(책등)
- detection성능 평가
- Web app(page2.1)

Support

- OCR 성능 평가
- modeling(제목)

멘토 이영완



퍼실 김상민
문준형

Q&A