

# 이미지 기반 영어 학습 보조 자료 제작 서비스

휴먼기계바이오공학부 황시은



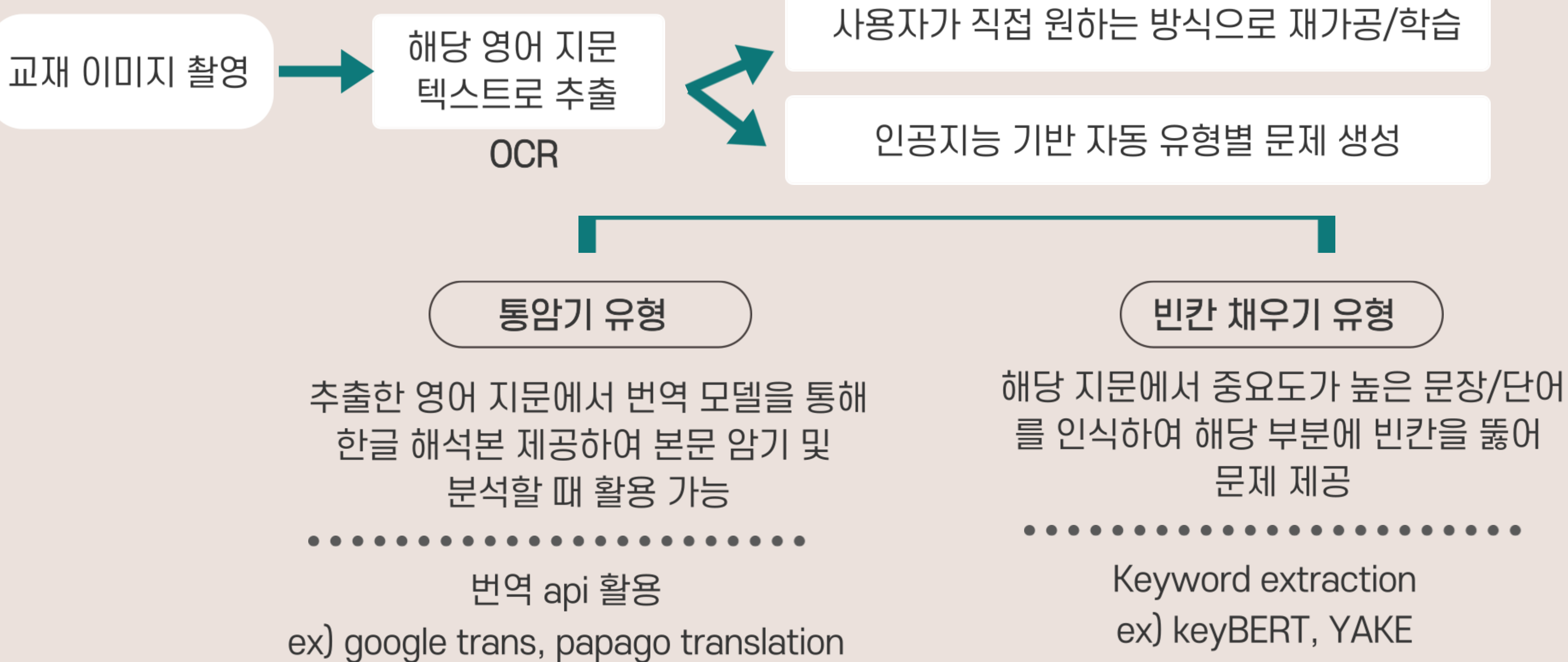
# 문제 정의

Main Problem: 교재 이미지 촬영만으로 영어 학습을 위한 보조 자료를 자동 생성해서 사용자의 효율적인 영어 학습을 위해 소요되는 시간과 노력을 줄인다.

## Sub problem

1. 긴 지문을 사용자가 직접 타이핑 할 필요 없이 이미지만 촬영하면 OCR을 통해 자동으로 텍스트로 추출하여 제공한다.
2. 추출된 지문의 해석본을 생성하여 사용자가 지문 분석 및 이해를 위해 사용할 수 있도록 한다.
3. 지문에서 중요도가 높은 단어/구절을 자동으로 빈칸 채우기 유형의 문제를 생성한다.

# AI 서비스 설계



# OCR 테스트 및 성능 비교

## Google Vision API

an astonishing array of tropical fish, the most remarkable of which are the five hundred of glaciers and severe drought, the incredible range of cichlid diversity in the lake must have developed in a relatively short evolutionary time period. In fact, sediment samples taken The world's second largest freshwater lake after Lake Superior, Lake Victoria is home to years ago at the end of the most recent ice age.

TOEFL 교재 이미지로 두 가지 방식을 활용하여  
OCR 성능 비교

--> Google Vision API의 성능이 확연하게 좋음

## 발생한 한계점들

- 교재에 단락 앞에 있는 단락 번호 주변에서 정확도 저하
- 지문 외에 문제도 같이 있는 페이지의 경우 문제와 지문을 구분할 수 있는 방법 필요

## Teserract

ic) A A>}.  
3 | Cichlids on  
freshwater lake after Lake Superior, Lake Victorig is  
t fre  
P om  
arkable of which are the ¢ 8 ty  
ishing array of tropical fish, the most rem in size from three cent five t  
- vane : f cichlids a group of fish ranging aa ina tied only af Meters Do  
different ET . eed more amazing when considering y sixty SPecigg 7  
meter. This variety  
inhabit all of Western Europe.  
{ The world's second larges  
ely dried up several times in its history due to the format  
n

# OCR 한계점 극복 방안

## 1) 정규 표현식 활용

: 이미지를 그대로 OCR을 하고 인식된 텍스트 상에서 정규 표현식을 활용하여 숫자 이후의 텍스트를 없애는 식으로 텍스트 상의 규칙성을 이용한다.

## 2) Document Understanding (Document Layout Detection)

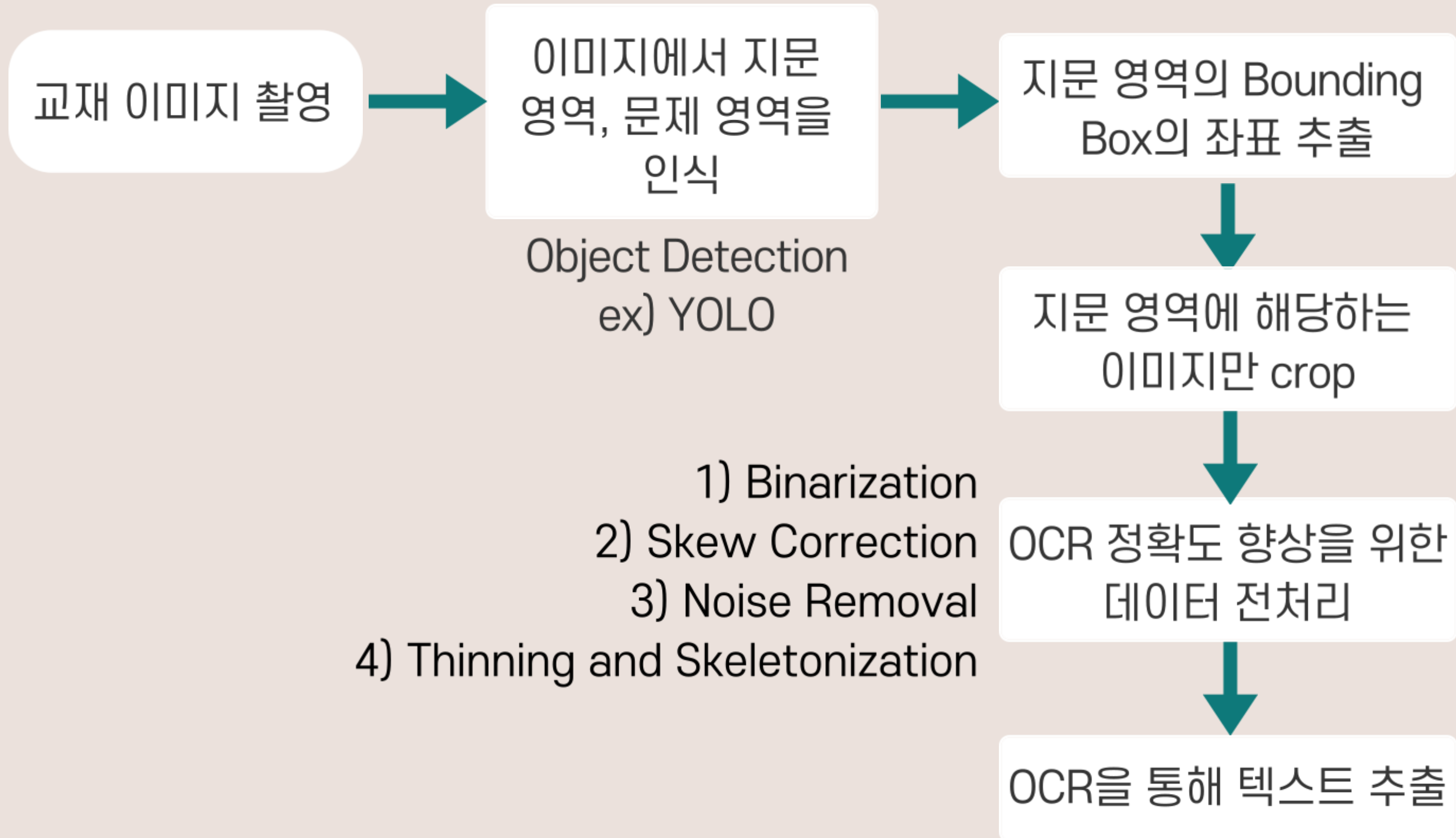
: 문서에서 필요한 데이터를 추출하여 유의미한 정보를 도출하는 것을 말한다. 영수증, 송장 등의 문서에서 원하는 정보를 얻기 위해 사용된다.

Ex) Chargrid, BERTgrid, CUTIE, LayoutLM, LayoutParser

## 3) Object Detection

: 객체 인식 모델을 직접 학습시켜서 지문 영역과 문제 영역을 각각 인식하도록 한다. 지문 영역에 대한 정보를 얻고 해당 영역에서만 텍스트를 읽어오는 방식으로 원하는 지문의 텍스트만 읽는다.

# OCR 한계점 극복 방안

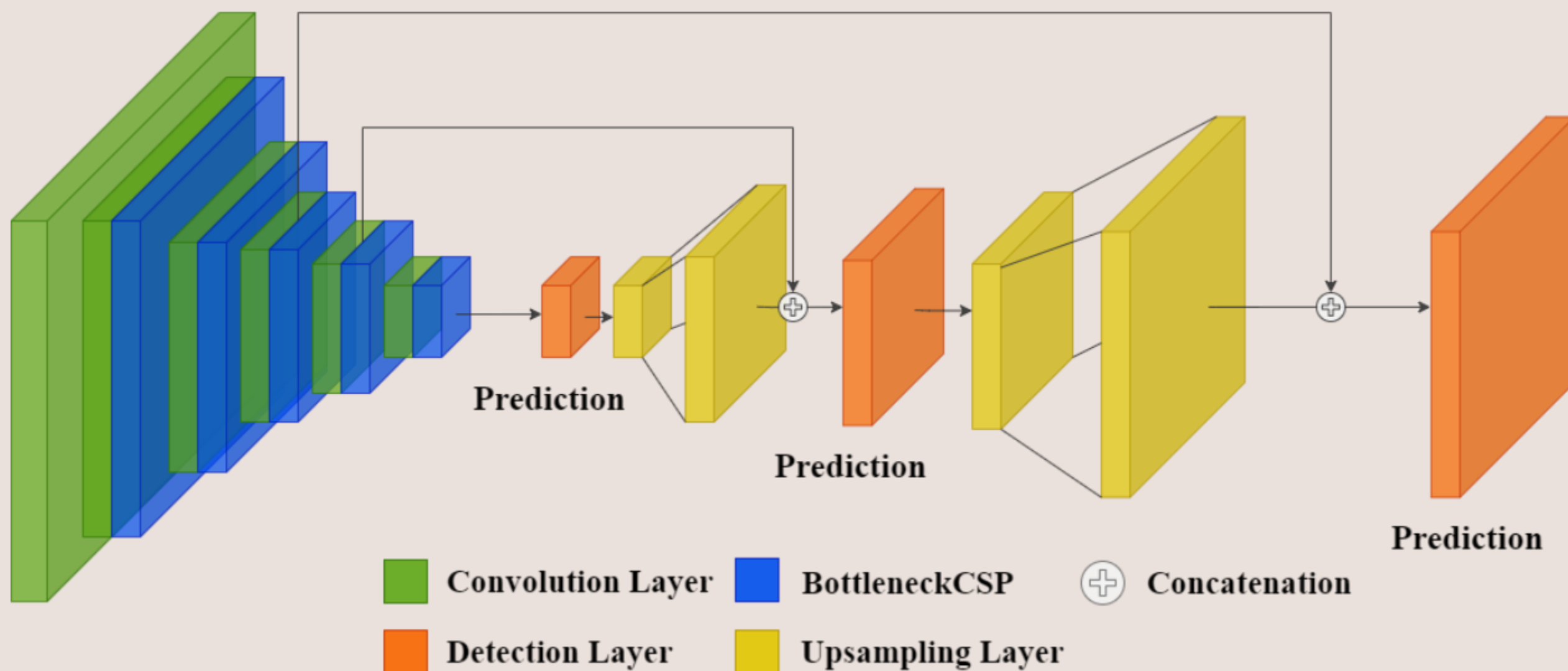




# OCR 한계점 극복 방안 - Object Detection

## YOLO v5

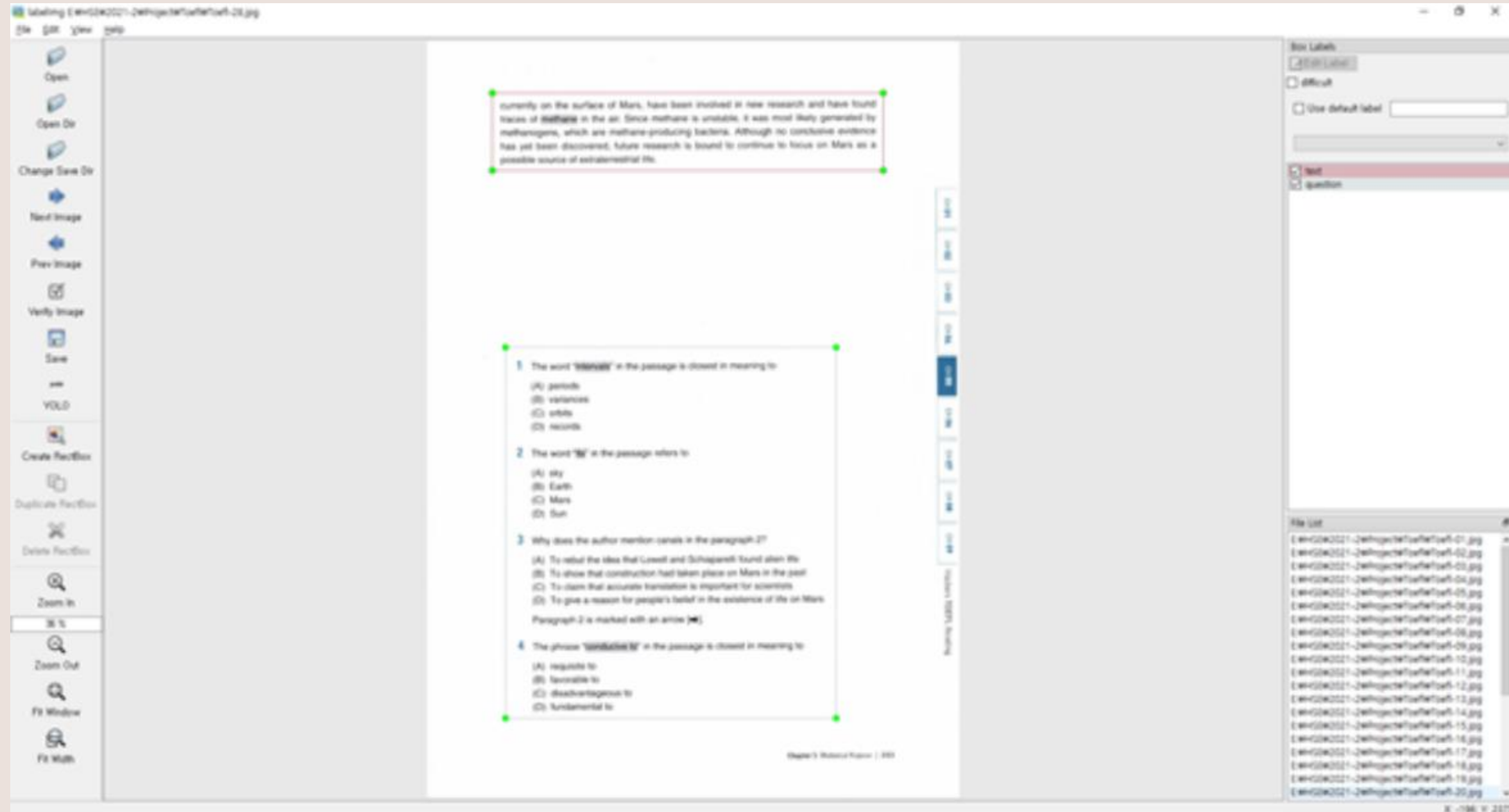
- YOLO v5 (You Only Look Once)는 대표적인 객체 탐지 모델이다.
- 이미지 전체를 단 한번의 스캔하는 one-stage algorithm이며, real-time object detection이 가능하여 매우 빠른 속도로 객체를 탐지하여 실시간으로 이용이 가능하다.



# YOLO V5

## Dataset

- labellmg 툴을 이용하여 직접 토플 교재를 라벨링 하였다.
- Text, question이라는 두 개의 라벨을 가지도록 yolo 포맷으로 라벨링했다.
- 총 32개의 이미지를 라벨링하였고, 22개를 training, 10개를 validation에 활용하였다.

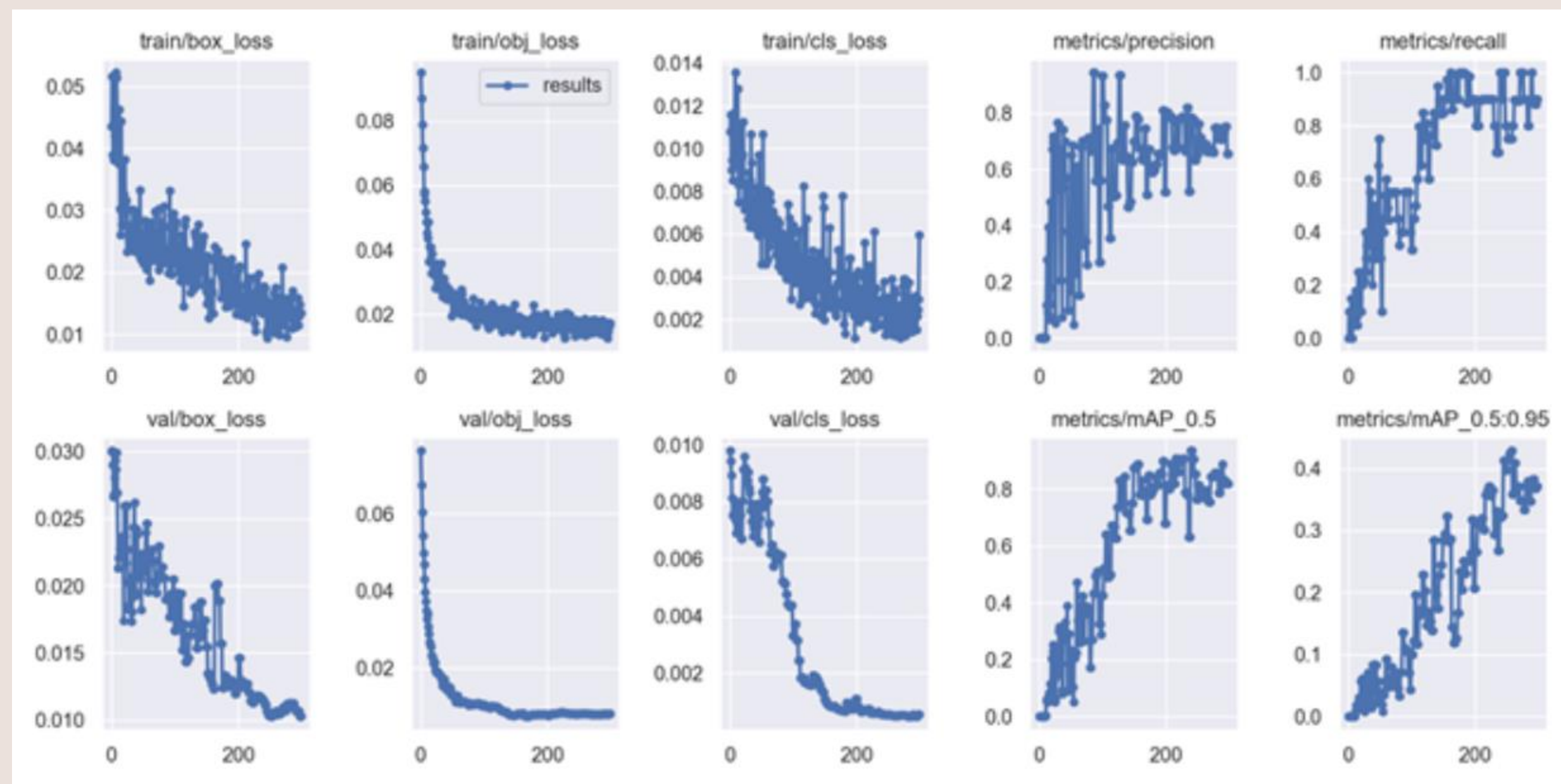




# YOLO V5

## Training

- 총 2개의 class (text, question)
- 이미지는 1280 x 1280으로 resizing, batch size 2, epoch 300으로 학습했다.
- Training 결과 loss함수가 잘 줄어드는 것을 확인했으며 최대 0.9351의 mAP를 얻었다.



# YOLO V5 – Results

## Best Case

Text 0.51

parenting and attempt multiple matings each season, sexual selection by females can play a powerful evolutionary role. When females have a wide variety of partners to choose from, they are allowed to be more particular about their mates. The males, on the other hand, have to find some way to overcome the enormous competition. In cichlids, it has led to spectacular male courtship exhibitions, using colors designed to entice females into spawning. Since females endemic to Lake Victoria select partners with particular features, it increases the chances that groups preferring alternate characteristics will then sympatrically divide into separate species.

- 6 Unfortunately, the incredible diversity has been decreasing and now over 250 indigenous types can no longer be found in Lake Victoria. In the 1950s, over the objections of wildlife ecologists, the British colonial government, hoping to give the local fishery a bigger fish to catch, released a non-native species, the Nile perch, into the lake. Commonly referred to as a voracious predator, the perch can grow into a hefty six-foot, two-hundred-pound giant. The thinking at the time was that the large predator would feed on the smaller fishes (i.e. the numerous small cichlids) and lead an immense growth in the fishing industry in the area. It was quick to establish itself and fundamentally changed the lake's ecology by eating, to the point of extinction, at least half of the cichlid species. At present, cichlids

- 1 The word "severe" in the passage is closest in meaning to
- (A) volatile  
(B) barren  
(C) harsh  
(D) arid
- 2 The word "deterred" in the passage is closest in meaning to
- (A) bemoaned  
(B) stopped  
(C) ordained  
(D) protracted
- 3 The word "colonize" in the passage is closest in meaning to
- (A) settle  
(B) desert  
(C) join with  
(D) conform to
- 4 According to paragraph 3, why did cichlids segregate into separate populations?
- (A) They faced intense opposition from other invasive species.  
(B) They were cut off from one another by fluctuating water levels in the lake.  
(C) Their growing numbers created destructive competition within the species.  
(D) Their initial habitat in the lake was unfavorable to survival.

Paragraph 3 is marked with an arrow [➡].

## Worst Case

lighting and costuming.

- 5 Gradually, Fuller's innovative techniques evolved into a performance she called "The Serpentine Dance," and in February 1892, at the revue "Uncle Celestin" in New York, she presented her free dance routine to considerable acclaim. The show traveled to Paris, receiving a warm reception that persuaded her to remain in France to become a regular entertainer at the Folies Bergere, the premiere international venue for experimental performances. Here, Fuller, the first modern dancer to perform in Europe, headlined numerous shows, most significantly "Fire Dance," in which she danced on glass illuminated from below. At a time when the Art Nouveau movement was all the rage, she burst upon the scene as its living embodiment and garnered the respect of many French artists such as Henri de Toulouse-Lautrec, who painted many of her promotional posters, and Auguste Rodin, who later used Fuller as a model for his sculptures. Scientists also flocked to Paris to witness firsthand the avant-garde lighting designs. By 1908, when she opened her own natural movement school, the European community, and, to some extent, her American compatriots, had accepted free dance, in which her inventions and talent coalesced as a serious art form.

### Glossary

stock companies a theatrical company that performs plays from a repertoire

Question 0.31

- 1 According to the passage, which of the following statements regarding the early career of Loie Fuller is true?
- (A) Fuller was inspired by the progressive approach utilized in the Wild West Show.  
(B) Fuller received tribute for her natural style of dance and her technical proficiency.  
(C) Fuller played her first important theatrical role in a production in New York City.  
(D) Fuller compensated for deficiencies in her education by studying innovative dancers.
- 2 The word "manipulating" in the passage is closest in meaning to
- (A) controlling  
(B) inspecting  
(C) layering  
(D) detaching
- 3 The word "splendor" in the passage is closest in meaning to
- (A) magnificence  
(B) transparency  
(C) vacancy  
(D) benevolence



# 지문 텍스트 추출

## 기존 이미지에 대한 Detection 결과

Hackers TOEFL READING

text 0.44

mills, glass factories, and metal foundries sprang up in clusters close to huge manufacturing plants. <sup>7</sup>At the same time, machine tool and die makers who provided all the precision parts and mechanical devices for the factory infrastructure flourished as technology continued to improve, the offshoots of which reared a range of new and exciting careers, particularly of interest to the large groups of young men returning from military duty around this time, and kept unemployment rates down while boosting overall spending. In the end, this booming economy together with better automation allowed Ford to drop his selling price from about \$850 in 1908 to less than \$300 in 1925, at which time sales figures skyrocketed as average customers could now afford to purchase a vehicle, feeding back into the overall health of companies that were involved with car making.

5 The mounting popularity of personal transportation ultimately had repercussions on the cities and countryside as they rapidly changed to make room for the astounding increase in cars. Dirt roads were replaced with paved streets and freeways while urban developers hurried to design gutters, curbs, traffic signs, and other infrastructure that was previously unnecessary. <sup>8</sup>Leisure time now circled around car ownership, bringing forth entertainment venues such as the drive-in theater or touring on newly constructed highways with accompanying gas stations, rest stops, billboards, and even motels to lodge weary travelers along the way. The entire cultural landscape was altered to focus on driving and the needs of the automobile consumer, the impact of which has had major reverberations on the economic survival of the country.

CH 01

CH 02

CH 03

CH 04

CH 05

CH 06

CH 07

CH 08

CH 09

Hackers TOEFL Reading

1 The word "rapid" in the passage is closest in meaning to

(A) fast

(B) strategic

(C) broad

(D) clever

2 Which of the following is true of the automobile industry in America in paragraph 1?

(A) At the same time, machine tool and die makers who provided all the precision parts and mechanical devices for the factory infrastructure flourished as technology continued to improve, the offshoots of which reared a range of new and exciting careers, particularly of interest to the large groups of young men returning from military duty around this time, and kept unemployment rates down while boosting overall spending.

(B) In the end, this booming economy together with better automation allowed Ford to drop his selling price from about \$850 in 1908 to less than \$300 in 1925, at which time sales figures skyrocketed as average customers could now afford to purchase a vehicle, feeding back into the overall health of companies that were involved with car making.

(C) The mounting popularity of personal transportation ultimately had repercussions on the cities and countryside as they rapidly changed to make room for the astounding increase in cars. Dirt roads were replaced with paved streets and freeways while urban developers hurried to design gutters, curbs, traffic signs, and other infrastructure that was previously unnecessary.

(D) Leisure time now circled around car ownership, bringing forth entertainment venues such as the drive-in theater or touring on newly constructed highways with accompanying gas stations, rest stops, billboards, and even motels to lodge weary travelers along the way. The entire cultural landscape was altered to focus on driving and the needs of the automobile consumer, the impact of which has had major reverberations on the economic survival of the country.

## Crop 후의 이미지

mills, glass factories, and metal foundries sprang up in clusters close to huge manufacturing plants. <sup>7</sup>At the same time, machine tool and die makers who provided all the precision parts and mechanical devices for the factory infrastructure flourished as technology continued to improve, the offshoots of which reared a range of new and exciting careers, particularly of interest to the large groups of young men returning from military duty around this time, and kept unemployment rates down while boosting overall spending. In the end, this booming economy together with better automation allowed Ford to drop his selling price from about \$850 in 1908 to less than \$300 in 1925, at which time sales figures skyrocketed as average customers could now afford to purchase a vehicle, feeding back into the overall health of companies that were involved with car making.

The mounting popularity of personal transportation ultimately had repercussions on the cities and countryside as they rapidly changed to make room for the astounding increase in cars. Dirt roads were replaced with paved streets and freeways while urban developers hurried to design gutters, curbs, traffic signs, and other infrastructure that was previously unnecessary. <sup>8</sup>Leisure time now circled around car ownership, bringing forth entertainment venues such as the drive-in theater or touring on newly constructed highways with accompanying gas stations, rest stops, billboards, and even motels to lodge weary travelers along the way. The entire cultural landscape was altered to focus on driving and the needs of the automobile consumer, the impact of which has had major reverberations on the economic survival of the country.

Prediction 결과에서 얻은 좌표를 이용하여 이미지를 자르고 OCR 적용

## Crop한 이미지에 대한 OCR 결과

converted(0.0, 0.496555, 0.737679, 0.413076) to (0.1277, 0.5416, 0.737679, 0.413076)  
crop to (241.0, 1.458e+03, 1632.9994215000002, 346.000068)  
Texts:  
mills, glass factories, and metal foundries sprang up in clusters close to huge manufacturing plants. 'At the same time, machine tool and die makers who provided all the precision parts and mechanical devices for the factory infrastructure flourished as technology continued to improve, the offshoots of which reared a range of new and exciting careers, particularly of interest to the large groups of young men returning from military duty around this time, and kept unemployment rates down while boosting overall spending. In the end, this booming economy together with better automation allowed Ford to drop his selling price from about \$850 in 1908 to less than \$300 in 1925, at which time sales figures skyrocketed as average customers could now afford to purchase a vehicle, feeding back into the overall health of companies that were involved with car making. The mounting popularity of personal transportation ultimately had repercussions on the cities and countryside as they rapidly changed to make room for the astounding increase in cars. Dirt roads were replaced with paved streets and freeways while urban developers hurried to design gutters, curbs, traffic signs, and other infrastructure that was previously



# 지문 텍스트 추출

## 통암기 유형

추출된 텍스트의 번역본을 생성하여 학생들이 본문 분석 및 암기 시험에 활용할 수 있는 자료를 생성한다.

--> Googletrans API 이용

<Text>

mills, glass factories, and metal foundries sprang up in clusters close to huge manufacturing plants. 'At the same time, machine tool and die makers who provided all the precision parts and mechanical devices for the factory infrastructure flourished as technology continued to improve, the offshoots of which reared a range of new and exciting careers, particularly of interest to the large groups of young men returning from military duty around this time, and kept unemployment rates down while boosting overall spending. In the end, this booming economy together with better automation allowed Ford to drop his selling price from about \$850 in 1908 to less than \$300 in 1925, at which time sales figures skyrocketed as average customers could now afford to purchase a vehicle, feeding back into the overall health of companies that were involved with car making. The mounting popularity of personal transportation ultimately had repercussions on the cities and countryside as they rapidly changed to make room for the astounding increase in cars. Dirt roads were replaced with paved streets and freeways while urban developers hurried to design gutters, curbs, traffic signs, and other infrastructure that was previously unnecessary. °Leisure time now circled around car ownership, bringing forth entertainment venues such as the drive-in theater or touring on newly constructed highways with accompanying gas stations, rest stops, billboards, and even motels to lodge weary travelers along the way. The entire cultural landscape was altered to focus on driving and the needs of the automobile consumer, the impact of which has had major reverberations on the economic survival of the country. |

문제 생성



<통암기 유형>

밀스, 유리 공장 및 금속 파운드리라는 거대한 제조 공장에 가까운 클러스터에서 위로 솟아갑니다. '동시에, 공장 인프라에 대한 모든 정밀 부품과 기계 장치를 제공 한 공장 인프라를 제공 한 공장 도구 및 다이 제조사는 새로운 개선과 흥미진진한 경력의 범위를 키우고, 특히 큰 관심을 갖|춘 오프 샷 이 시간 동안 군사 의무에서 돌아 오는 젊은이들의 그룹은 전체 지출을 향상시키면서 실업률을 유지했습니다. 결국, 더 나은 자동화와 함께이 호황을 누리면서 1925 년에 1908 달러의 판매 가격을 약 300 달러 미만으로 판매 할 수있게 해주는 시간 판매 수치가 보통 고객이 차량을 구매할 수 있으므로 자동차 제조에 관여 한 기업의 전반적인 건강. 개인 교통의 장착 인기는 궁극적으로 자동차의 놀라운 증가를위한 공간을 만들기 위해 급속하게 변화했기 때문에 도시와 시골에서의 반향이있었습니다. 비포장 도로는 포장 된 거리와 고속도로로 교체되었습니다. 도시 개발자들은 이전에 불필요한 거터, 연석, 교통 표지 및 기타 인프라를 설계하기 위해 서둘렀습니다. ° 자동차 소유권 주위에 휴가 시간을 모으고, 주유소와 같은 새로 건설 된 고속도로와 같은 새로 건설 된 고속도로와 같은 엔터테인먼트 장소를 열어주는 가스 스테이션, 휴식 중단, 빌보드 및 심지어 도중에 지친 여행자를 덜어줍니다. 전체 문화 경관은 자동차 소비자의 운전 및 필요에 집중하는 데 집중하여 이국의 경제적 생존에 대한 주요 잔향이있는 영향을 받았습니다.

# 문제 생성

## 빈칸 채우기 유형

: Keyword extraction 모델을 활용하여 중요도가 높은 단어/구절들을 선별하고 해당 부분에 빈칸을 뚫어 학생들이 직접 빈칸을 채우며 공부할 수 있는 문제와 답변을 제공한다.

### [ KeyBERT ]

1. input 문서가 pre-trained BERT 모델을 이용하여 일정한 크기의 벡터로 embedded된다.
2. 동일 문서에서 tf-idf vectorizer나 count vectorizer를 이용하여 키워드를 추출한다.
3. 추출된 키워드는 위에서 document embedding에 사용한 모델로 동일하게 일정 크기의 벡터로 embedded 된다.
4. 키워드와 문서는 각각 같은 공간에 represented 되었으며 이때 이들 간의 cosine similarity를 계산한다.
5. 추출된 similarity를 기반으로 가장 관련도가 높은 키워드부터 정렬한다.



# 문제 생성

## 빈칸 채우기 유형

<빈칸 유형>

mills, glass factories, and metal foundries sprang up in clusters close to huge manufacturing plants. 'At the same time, machine tool and die makers who provided all the precision parts and mechanical devices for the \_\_\_\_\_ infrastructure flourished as technology continued to improve, the offshoots of which reared a range of new and exciting careers, particularly of interest to the large groups of young men returning from military duty around this time, and kept unemployment rates down while boosting overall spending. In the end, this booming \_\_\_\_\_ together with better automation allowed Ford to drop his selling price from about \$850 in 1908 to less than \$300 in 1925, at which time sales figures skyrocketed as average customers could now af\_\_\_\_\_ to purchase a vehicle, feeding back into the overall health of companies that were involved with car making. The mounting popularity of personal transportation ultimately had repercussions on the cities and countryside as they rapidly changed to make room for the astounding increase in \_\_\_\_\_. Dirt roads were replaced with paved streets and freeways while urban developers hurried to design gutters, curbs, traffic signs, and other infrastructure that was previously unnecessary. °Leisure time now circled around car ownership, bringing forth entertainment venues such as the drive-in theater or touring on newly constructed highways with accompanying gas stations, rest stops, billboards, and even motels to lodge weary travelers along the way. The entire cultural landscape was altered to focus on driving and the needs of the \_\_\_\_\_ consumer, the impact of which has had major reverberations on the economic survival of the country.

[Answer]

Question 1: ford

Question 2: automobile

Question 3: economy

Question 4: factory

Question 5: cars

# Discussion

## 어려움 및 한계점

- 이미지를 직접 라벨링해서 너무 적은 숫자의 데이터셋으로 진행하였다. 더 많은 이미지를 사용한다면 더 좋은 결과를 얻을 수 있을 것이다.
- 번역 api의 성능 한계상 부자연스럽게 해석된 부분들이 존재했다.

## 개선방안

- 하나의 문제집에 대해서만 진행했기 때문에 다른 문제집의 경우 형식이나 포맷이 달라지기 때문에 범용적으로 사용할 수 있도록 하는 방식에 대해 고려해야 한다.
- 통암기나 빈칸 외에도 요약문 유형 등 좀 더 다양한 유형의 문제를 함께 제공하도록 발전시킬 수 있을 것이다.



**감사합니다**