

# 시각 장애인을 위한 편의점 음료 인식 프로젝트



이연경

박지예

sharpay62@sju.ac.kr / judy.greedy@gmail.com



# 주제 선정 배경

: 시각 장애인들은 혼자 편의점을 이용하기 어려운 불편함을 겪고 있다.



<https://www.youtube.com/watch?v=fq5xQaWaMOC>



<https://www.youtube.com/watch?v=woiF52DQHc4>



# 주제 선정 배경

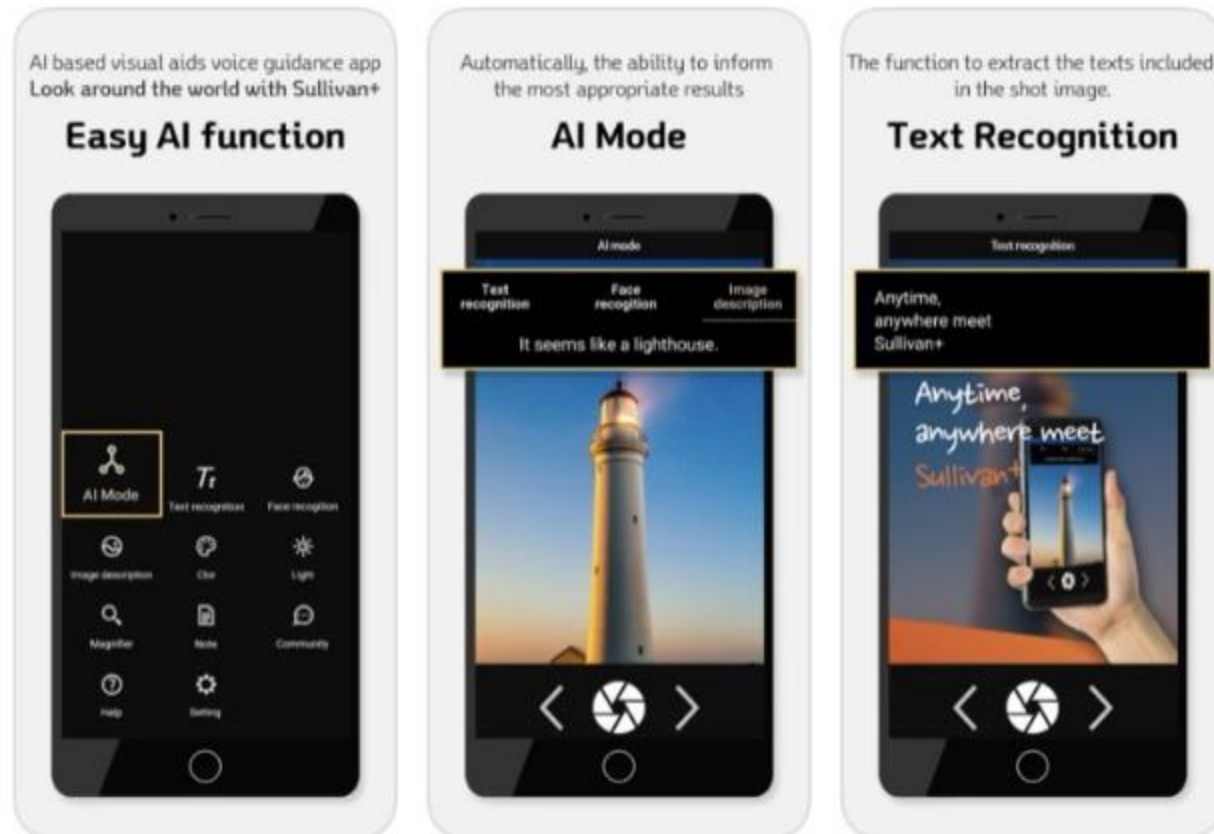
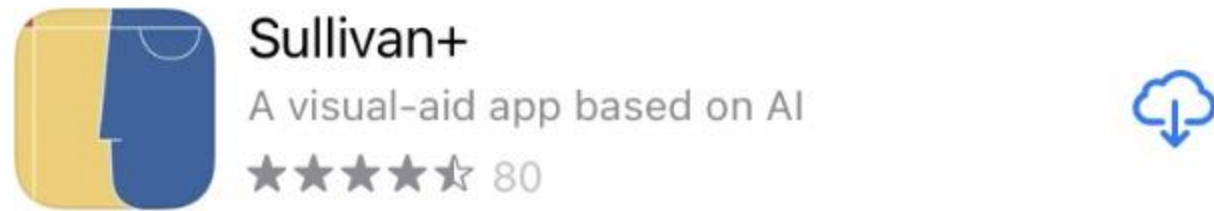
: 제품에 **점자가 없는 경우**가 대부분이고, 있는 경우에도 단순히 '음료'라고 적혀 있어 **정확한 종류를 파악하기 어려운 상태**





# 주제 선정 배경

: 카메라를 비추면 상황 설명을 해주는 ‘설리번’ 어플이 있기는 하지만,  
제품에 있는 모든 글자를 다 읽어줘서 어떤 상품인지 쉽게 알 수 없는 문제점 발생



스팸이랑, 김치? 뭐 이렇게 읽어주고 있습니다

# Dataset 구축하기

## -image crawling



캔음료 오리지널 코카콜라!

# Dataset 구축하기

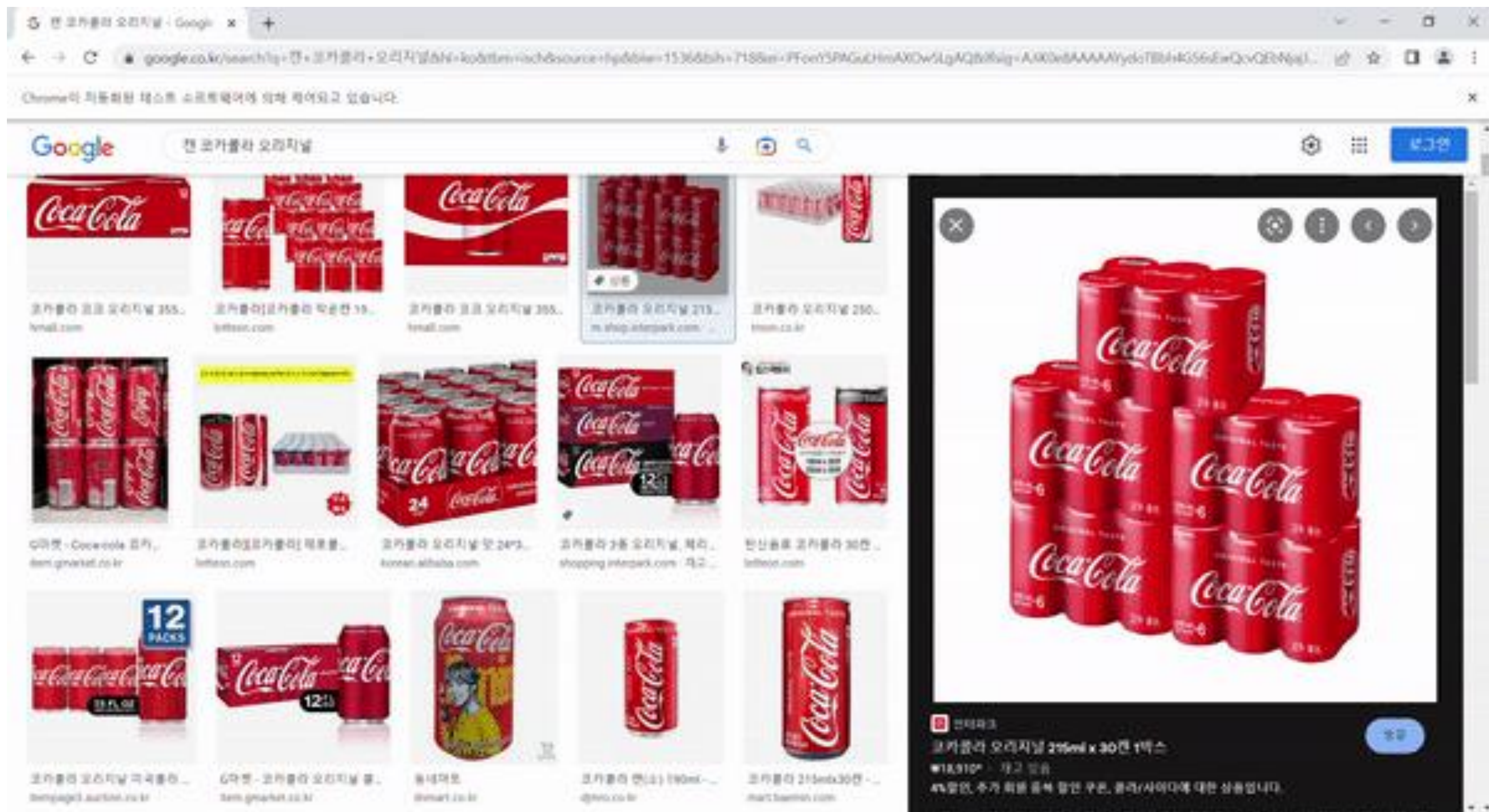
## -image crawling





# Dataset 구축하기

## -image crawling





# Dataset 구축하기

## -image crawling

: 검색 결과 오류로 인해 **다른 제품 사진이 저장**되거나

**여러 음료가 함께 있는 사진**은 학습에 혼동을 줄 수 있으므로 **삭제**함





# Dataset 구축하기

## -Dataset



Training Set - 453

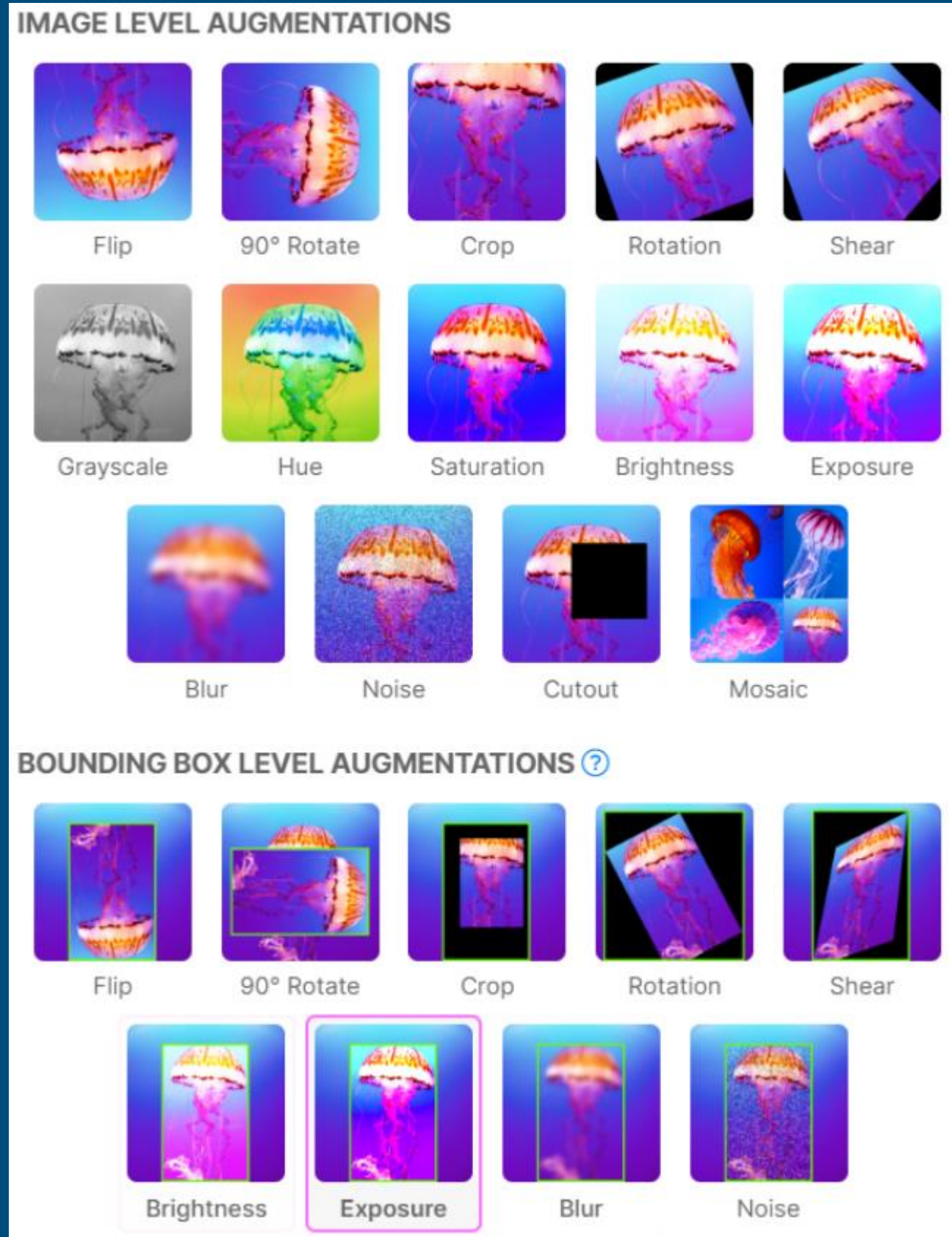
Validation Set - 42

Testing Set - 24



# Dataset 구축하기

## -image augmentation



- 음료를 잡을 때 손에 조금 가려질 경우

- cutout

- 거꾸로 들거나 돌려서 들 경우

- flip, rotation

- 카메라에 지나치게 가깝게 갖다대어

- 상품이 잘려서 보일 경우

- crop



# Dataset 구축하기

## -Dataset



Nano

YOLOv5n

4 MB<sub>FP16</sub>  
6.3 ms<sub>V100</sub>  
28.4 mAP<sub>COCO</sub>



Small

YOLOv5s

14 MB<sub>FP16</sub>  
6.4 ms<sub>V100</sub>  
37.2 mAP<sub>COCO</sub>



Medium

YOLOv5m

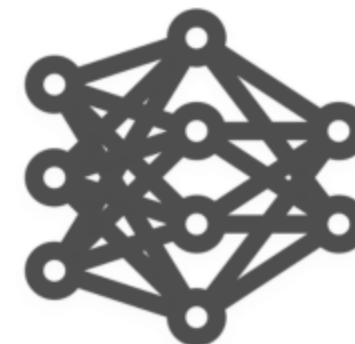
41 MB<sub>FP16</sub>  
8.2 ms<sub>V100</sub>  
45.2 mAP<sub>COCO</sub>



Large

YOLOv5l

89 MB<sub>FP16</sub>  
10.1 ms<sub>V100</sub>  
48.8 mAP<sub>COCO</sub>



XLarge

YOLOv5x

166 MB<sub>FP16</sub>  
12.1 ms<sub>V100</sub>  
50.7 mAP<sub>COCO</sub>

YOLOv5X 사용

Image : 320 batch : 16 epoch : 30



# 코카콜라가 있을 때

## ▼ 코카콜라가 있을 때

```
[ ] !python /content/yolov5/detect.py #  
--weights /content/yolov5/runs/train/result_cocacola/weights/best.pt #  
--img 320 #  
--conf 0.5 #  
--source /content/yolov5/runs/detect/exp/80_jpg.rf.a5b715074c69b690c03db4c99ccc2628.jpg
```

**detect:** weights=['/content/yolov5/runs/train/result\_cocacola/weights/best.pt'], source=/content/yolov5/runs/detect/exp/80\_jpg.rf.a5b715074c69b690c03db4c99ccc:  
YOLOv5 🚀 v6.2-61-gffbce38 Python-3.7.13 torch-1.12.1+cu113 CUDA:0 (Tesla P100-PCIE-16GB, 16281MiB)

Fusing layers...

YOLOv5x summary: 444 layers, 86173414 parameters, 0 gradients, 203.8 GFLOPs

image 1/1 /content/yolov5/runs/detect/exp/80\_jpg.rf.a5b715074c69b690c03db4c99ccc2628.jpg: 320x320 (no detections), 18.7ms

Speed: 0.3ms pre-process, 18.7ms inference, 0.3ms NMS per image at shape (1, 3, 320, 320)

Results saved to ../../../../runs/detect/exp13

이미지와 label을 같이 저장



# 코카콜라가 있을 때

Labels 폴더에 파일이 있을 때 '코카콜라가 있습니다'

```
import speech_recognition as sr

from gtts import gTTS

import os

import time

import playsound

def speak1(filedir, text):

    tts = gTTS(text=text, lang='ko')

    filename='voice_existence.mp3'

    tts.save(filedir + filename)

    #playsound.playsound(filename)

def speak2(filedir, text):

    tts = gTTS(text=text, lang='ko')

    filename='voice_noexistence.mp3'

    tts.save(filedir + filename)

    #playsound.playsound(filename)

filedir = '/content/drive/MyDrive/Team_Daiv/Convenience_store/results_custom_cocacola/exp14/' #exp넘버는 직접 설정하기

if not os.listdir(filedir + 'labels') :

    speak2(filedir, "Coca cola가 없습니다.")
else:
    speak1(filedir, "Coca cola가 있습니다.")
```



# 코카콜라가 있을 때





# 코카콜라가 없을 때

Labels 폴더에 파일이 없을 때 '코카콜라가 없습니다'

## ▼ 코카콜라가 없을 때

```
[ ] !python /content/yolov5/detect.py #  
--weights /content/yolov5/runs/train/result_cocacola/weights/best.pt #  
--img 320 #  
--conf 0.5 #  
--source /content/yolov5/runs/detect/exp/80_jpg.rf.a5b715074c69b690c03db4c99ccc2628.jpg
```

**detect:** weights=['/content/yolov5/runs/train/result\_cocacola/weights/best.pt'], source=/content/yolov5/runs/detect/exp/80\_jpg.rf.a5b715074c69b690c03db4c99ccc2628.jpg, data=../../../../../data/coco128  
YOLOv5 🚀 v6.2-61-gffbce38 Python-3.7.13 torch-1.12.1+cu113 CUDA:0 (Tesla P100-PCIE-16GB, 16281MiB)

Fusing layers...

YOLOv5x summary: 444 layers, 86173414 parameters, 0 gradients, 203.8 GFLOPs

image 1/1 /content/yolov5/runs/detect/exp/80\_jpg.rf.a5b715074c69b690c03db4c99ccc2628.jpg: 320x320 (no detections), 23.9ms

Speed: 0.3ms pre-process, 23.9ms inference, 0.5ms NMS per image at shape (1, 3, 320, 320)

Results saved to ../../../../../runs/detect/exp16

이미지와 label을 같이 저장



코카콜라가 있을 때





## 프로젝트의 한계

## 음료수의 다양성 부족

## 프로젝트 개발 시간적 한계

음료수의 종류를 줄여 Dataset 구축

## 낮은 정확성

제로콜라, 외관이 비슷한 다른 음료 등을

상세하게 구분할수 없음

## 프로젝트 구현 지연

프로젝트의 개발이 지연되어,

어플리케이션 구현은 개발 중



감사합니다 :D

Q&A