HUST

ĐẠI HỌC BÁCH KHOA HÀ NỘI

HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

ONE LOVE. ONE FUTURE.

ĐẠI HỌC
BÁCH KHOA HÀ NỘI
HANOI UNIVERSITY
OF SCIENCE AND TECHNOLOGY

Mini project report

Group 4 : Đỗ Tuấn Minh , Vũ Hải Đăng

ONE LOVE. ONE FUTURE.

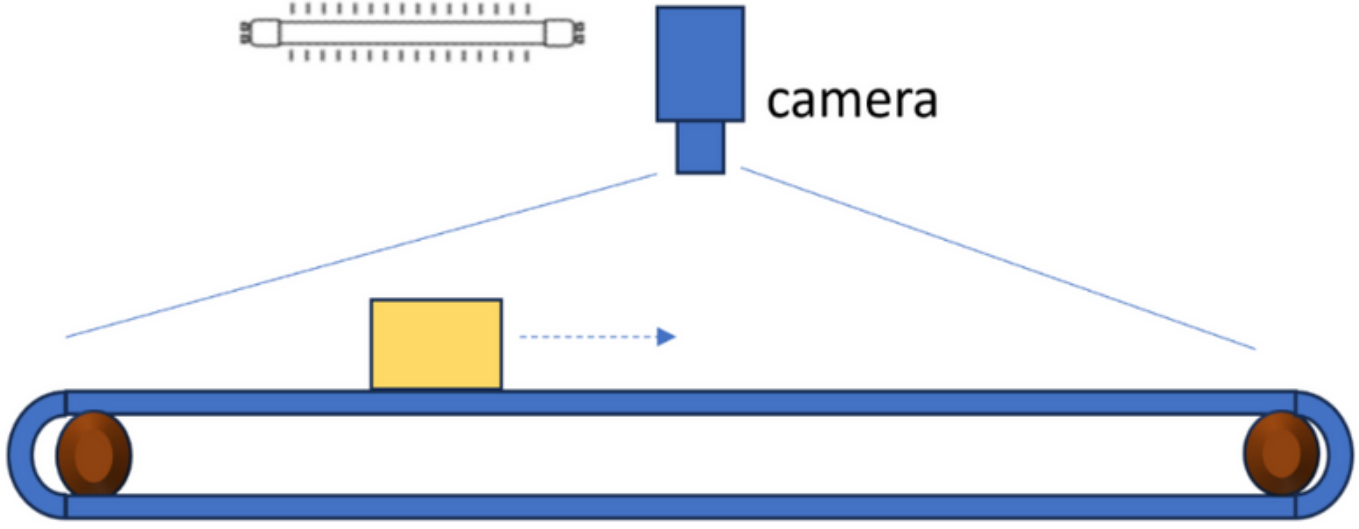
OUTLINE

1. Problem
2. Method
3. Details
4. Test and evaluate
5. Conclusion
6. Improvement

1. Problem

Problem statement :

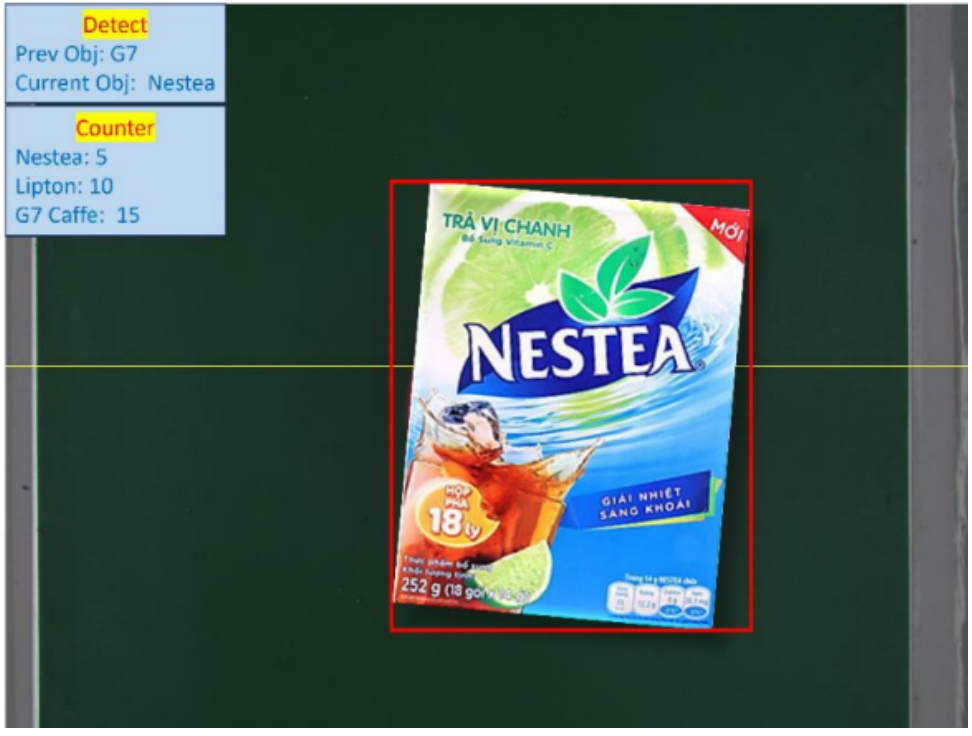
- Conveyor transport products
- Fixed cameras and light source
- Type of product : confectionery



1. Problem

Target :

1. Classify the goods
2. Count the number of each class
3. Show results



1. Problem

Challenges :

- Item size diversity
- Noise in background
- Limited resource
- Processing time

1. Problem

Good points :

- Item straight orbit
- Low moving speed
- Light conditions
- Single object per frame
- Fixed background

2. Method

How to select model ?

Survey of recent research on the **performance** and **inference speed** of object detection models

Model	Recall (%)	mAP (%)
CenterNet	58.0	81.2
YOLOv4	74.0	76.2
YOLOv5	79.4	84.1

[1]

Method	Params (MB)	inference time (ms)
Yolov5	89.0	6.6
CenterNet	74.99	28

[2]

Pipeline inference

Resize
CvtColor

Check object pass the marked zone

Image processing

Prediction

Bounding boxes infor

Fine-tuned Yolov5n

Create counter

Show image ouput

Post processing

DATA PROCESSING

- After being resized, check whether the object has passed the marked area

Gray scale

Binary

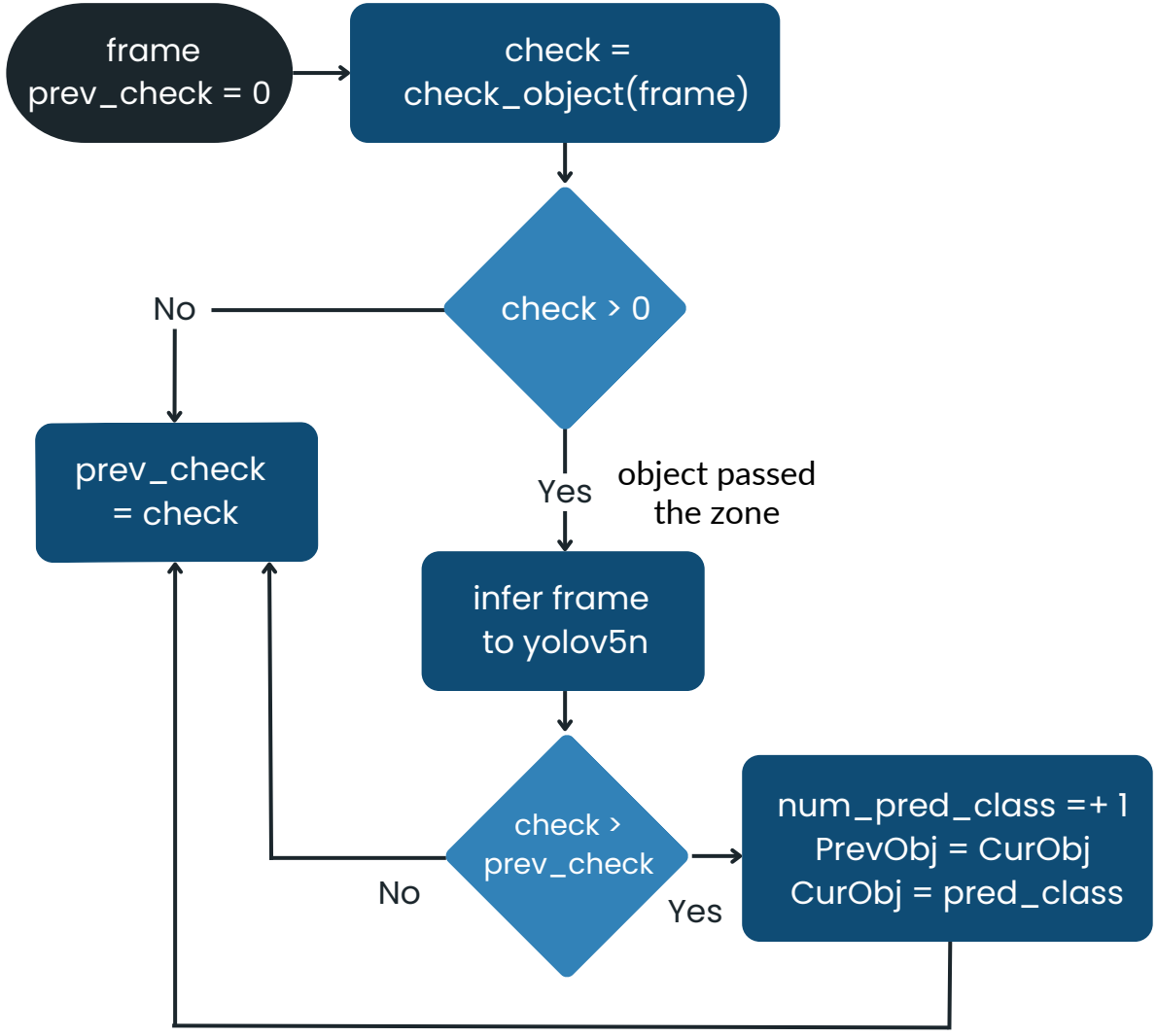
Median Blur

3. Details

COUNTER

```
table_data = [
    ["orion", 0],
    ["tipo", 0],
    ["y40", 0],
    ["nestea_hoaqua", 0],
    ["chanh", 0],
    ["nestea_atiso", 0],
    ["karo", 0],
    ["jack-jill", 0],
    ["g7", 0],
    ["dilmah", 0]
]
```

```
detection = [
    ["Prev", ""],
    ["Current", ""],
]
```



3. Details

Training

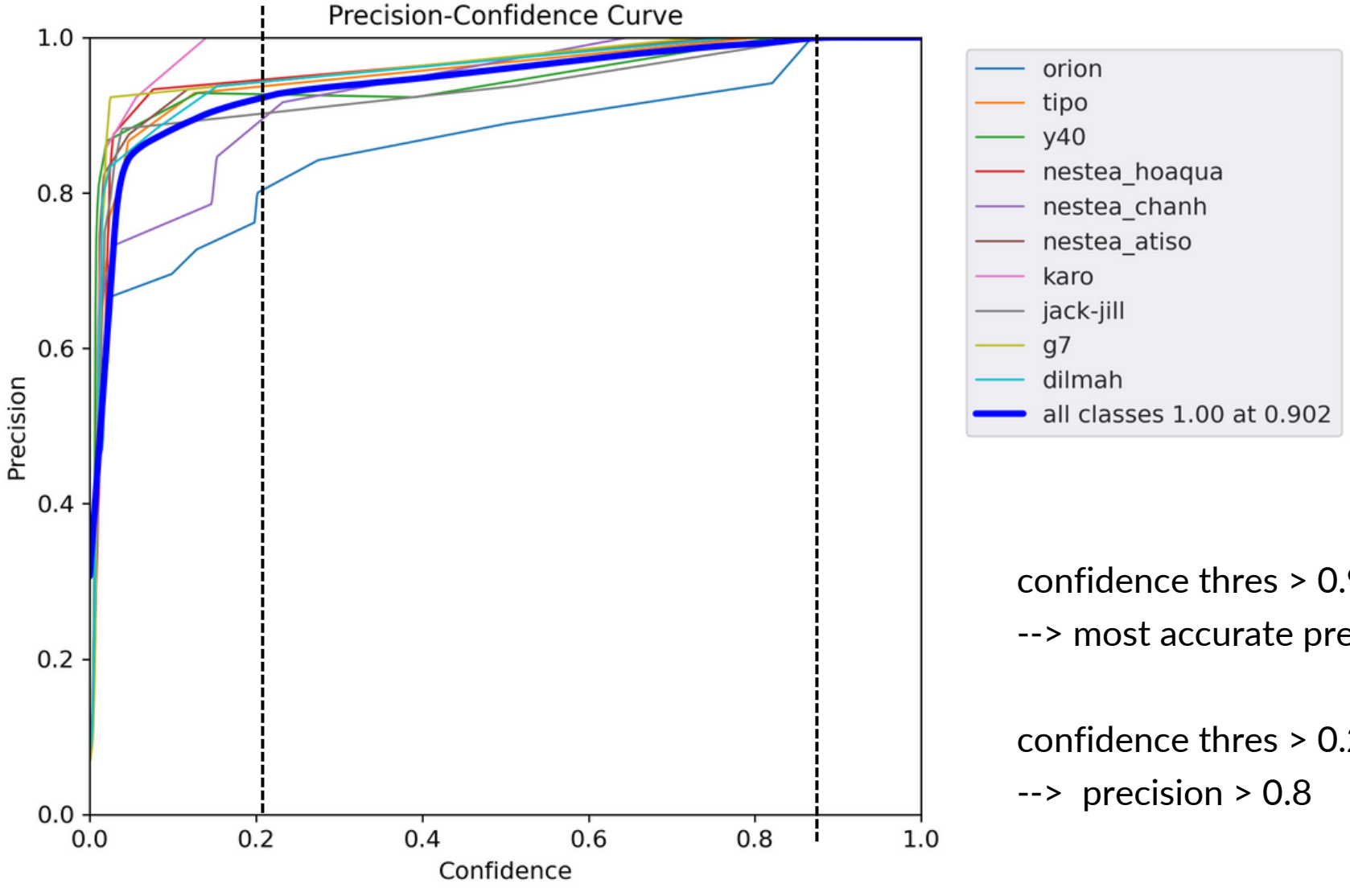
Data :

- cellphone_video_for_train (all to train data)
- conveyor_video_for_train (similar to test video)
- train/val/test = 1177 / 151 / 138
- classes ratio range from 7 - 11%

Labels :

- using tool LabelImg then trained YOLOv5 for auto labelling

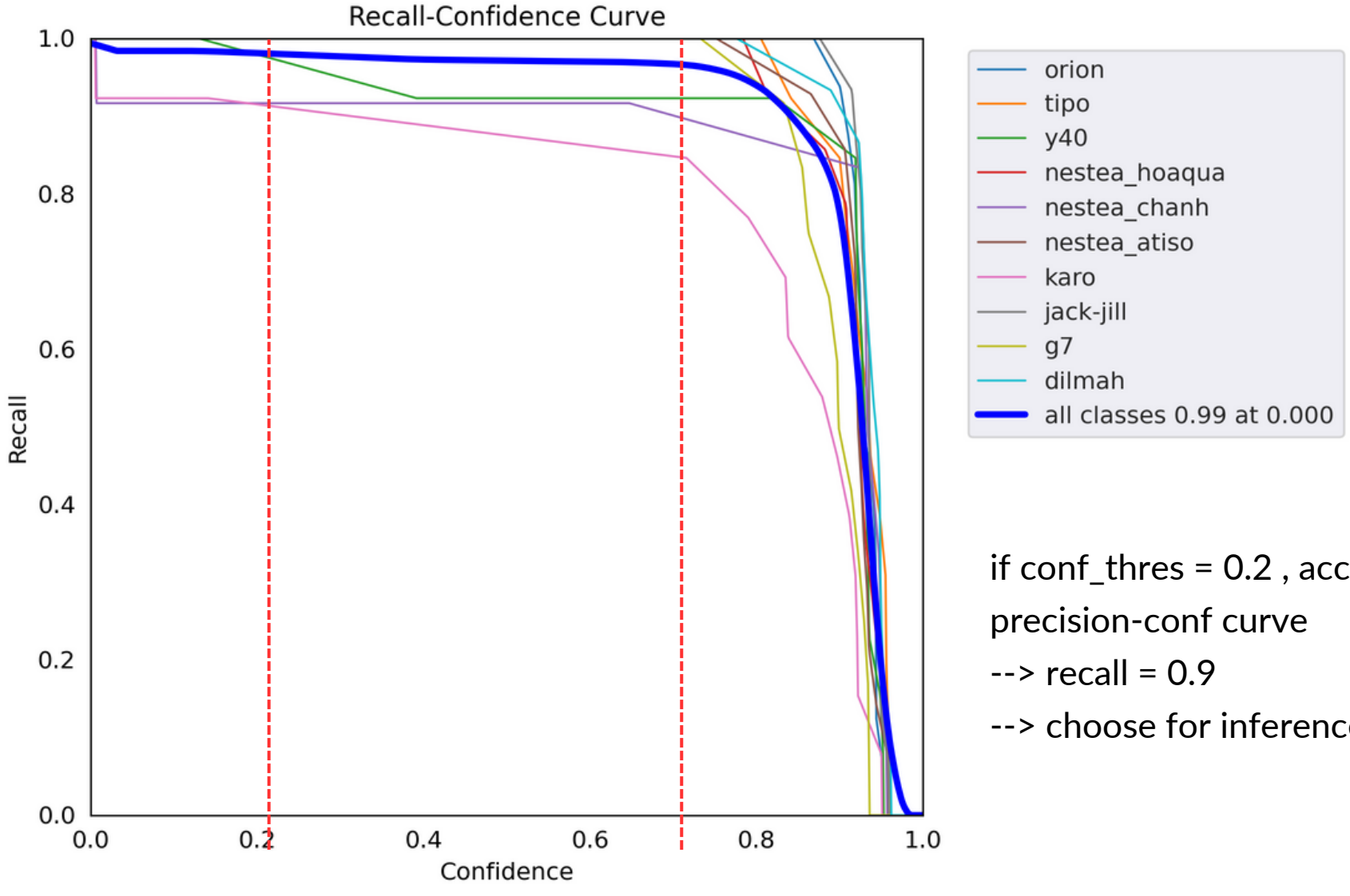
4. Test and Evaluate



confidence thres > 0.902
--> most accurate prediction

confidence thres > 0.2
--> precision > 0.8

4. Test and Evaluate



if conf_thres = 0.2 , according to precision-conf curve

--> recall = 0.9

--> choose for inference thres

4. Test and Evaluate

Inference

Video inference : 3m14s length and 16 fps

Hardware :

- CPU 11th Gen Intel(R) Core(TM) i7-11800H @ 2.30GHz
- 8 cores, 12GB RAM

Inference time

- frame no object : 4-6 ms
- frame with object : 35-50ms
- whole video : 80 seconds

4. Test and Evaluate

Inference

Output screenshot

5.Conclusion

- With small training data, the counter return results accurately but the bounding boxes do not overfit
- **Satisfy** the 3 requirements of project output
- **Limitation:** when having new item, must label manually again and then train the yolov5 again

6.Improvement

Improve proposed method :

Only infer the frame that object just passed the marked zone, then save that frame to .png at local to check results

--> only have to infer 11 frames

--> inference time :

- frame no object : ~1 ms
- frame object : 70-80 ms
- video inference time : 20s

6.Improvement

New approach :

- Image processing to detect boxes, the crop
- Process data to create training dataset
- Train simple CNN to classify

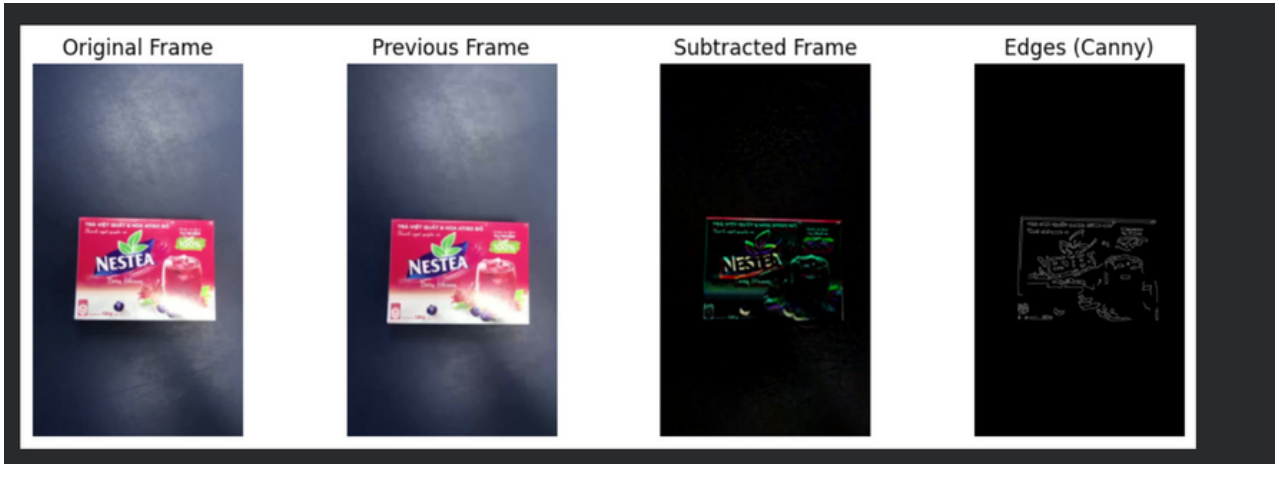
--> Benefit :

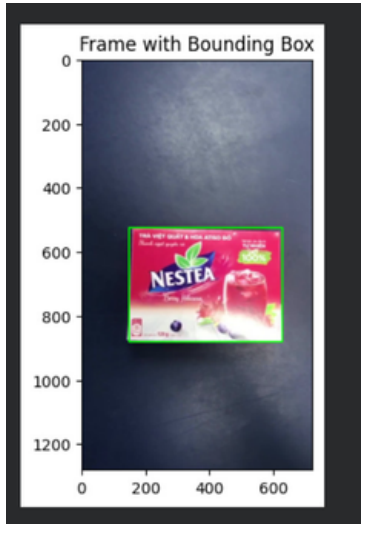
- Auto generate bounding boxes, training data
- Maybe lighter

Key point of this method : Image processing techniques to detect boxes

6.Improvement

```python
def create_bounding_box(frame, pre_frame):

    subtracted = cv2.subtract(frame, pre_frame)

    gray = cv2.cvtColor(subtracted, cv2.COLOR_BGR2GRAY)

    blurred_image = cv2.medianBlur(gray,9,0)

    min_threshold = 20
    max_threshold = 80
    edges_subtract = cv2.Canny(blurred_image, min_threshold, max_threshold)

    points = list(np.where(edges_subtract > 0))

    cv2.rectangle(frame, (min(points[1]), min(points[0])), (max(points[1]), max(points[0])), (0, 0, 255), 3)

    return frame
```





6.Improvement

- Object Tracking (Yolov5 + DEEPSORT, SORT)
- Quantization
- Ensemble Learning

Final words

- Learn to use, survey papers about CNN models : CenterNet, YOLO, ...
- Get familiar with the process of developing AI, object detection project end-to-end
- Teamwork ability, supporting each other

HUST

hust.edu.vn fb.com/dhbkhn

THANK YOU !