Piano Keyboard Detection Experiments Aug 16th~18th, 2023

Index_

- 1. Introduction of dataset for experiment
- 2. Evaluation Metric for Object Detection (YOLOv5, SSD-MobileNet)
- 3. Comparison the results

1. Introduction of dataset for experiment

1. Introduction of dataset for experiment

- There are 2,605 images from divided in 5 technique related videos.
- First, I split dataset 8:1:1 for train, test, and val respectively using "train_val_test_split.py" founded by searching on GitHub.

· Next, I downloaded these data files in my computer. Then, these are moved to YOLOv5 folder so that train the model for same data.



2. Evaluation Metric for Object Detection

- (YOLOv5, SSD-MobileNet)

2. Guidance on experimental procedures

YOLOv5

- · I obtained an evaluation metric for object detection according to IoU threshold using val.py provided.
- Used IoU threshold are 0.5, 0.6, 0.7, 0.8, 0.9 and 0.95.

SSD-MobileNet

- I obtained an evaluation metric for object detection according to IoU threshold using calculate_map_cartucho.py founded by searching on GitHub.
 - Used IoU threshold are from 0.5 to 0.95 in increments of 0.05.

2. Evaluation Metric for Object Detection

- YOLOv5

The image below is the command to get YOLOv5 model's evalutaion metric in cmd.

```
(base) C:\Users\User\Desktop\fingers detection\comparison_result_2023-08-14\Y0L0v5>python ./yolov 5/val.py --task "test" --data ./test.yaml --weights ./yolov5/runs/train/piano_yolov5s_results/weights/best.pt --device 0 --save-conf --iou-thres 0.5 val: data=./test.yaml, weights=['./yolov5/runs/train/piano_yolov5s_results/weights/best.pt'], bat ch_size=32, imgsz=640, conf_thres=0.001, iou thres=0.5, max_det=300, task=test, device=0, workers =8, single_cls=False, augment=False, verbose=False, save_txt=False, save_hybrid=False, save_conf= True, save_json=False, project=yolov5\runs\val, name=exp, exist_ok=False, half=False, dnn=False fatal: cannot change to 'C:\User\Desktop\fingers': No such file or directory Y0L0v5 2023-8-3 Python-3.10.11 torch-2.0.1 CUDA:0 (NVIDIA GeForce RTX 4080, 16376MiB)
```

The image below is the result of the YOLOv5 model's evalutaion metric.

```
YOLOv5s summary: 157 layers, 7012822 parameters, 0 gradients, 15.8 GFLOPs
test: Scanning C:\User\User\Desktop\fingers detection\comparison_result_2023-08-14\YOLOv5\keybo
Class Images Instances P R mAP50 mAP50-95: 100%|

all 261 261 1 1 0.995 0.819
Speed: 0.2ms pre-process, 3.6ms inference, 1.8ms NMS per image at shape (32, 3, 640, 640)
```

Precision: 1

• Recall: 1

• mAP50: 0.995

The image below is the command to get YOLOv5 model's evalutaion metric in cmd.

```
(base) C:\Users\User\Desktop\fingers detection\comparison_result_2023-08-14\Y0L0v5>python ./yolov 5/val.py --task "test" --data ./test.yaml --weights ./yolov5/runs/train/piano_yolov5s_results/weights/best.pt --device 0 --save-txt wal: data=./test.yaml, weights=['./yolov5/runs/train/piano_yolov5s_results/weights/best.pt'], bat ch_size=32, imgsz=640, conf_thres=0.001, iou thres=0.6, max_det=300, task=test, device=0, workers =8, single_cls=False, augment=False, verbose=False, save_txt=True, save_hybrid=False, save_conf=F alse, save_json=False, project=yolov5\runs\val, name=exp, esist_ok=False, half=False, dnn=False fatal: cannot change to 'C:\Users\User\Desktop\fingers': No such file or directory Y0L0v5 2023-8-3 Python-3.10.11 torch-2.0.1 CUDA:0 (NVIDIA GeForce RTX 4080, 16376MiB)
```

The image below is the result of the YOLOv5 model's evalutaion metric.

```
\text{YOLOV5s summary: 157 Layers, 7012822 parameters, 0 gradients, 15.8 GFLOPs test: Scanning C:\Users\User\Desktop\fingers detection\comparison_result_2023-08-14\YOLOV5\keybo test: New cache created: C:\Users\User\Desktop\fingers detection\comparison_result_2023-08-14\YOLOV5\keyboard\labels\test.cache

Class Images Instances P R mAP50 mAP50-95: 100%|

all 261 1 1 0.995 0.817

Speed: 0.2ms pre-process, 2.2ms inference, 1.0ms NMS per image at shape (32, 3, 640, 640)
```

Precision: 1

Recall: 1

mAP50: 0.995

The image below is the command to get YOLOv5 model's evalutaion metric in cmd.

```
(base) C:\Users\User\Desktop\fingers detection\comparison_result_2023-08-14\Y0L0v5>python ./yolov 5/val.py --task "test" --data ./test.yaml --weights ./yolov5/runs/train/piano_yolov5s_results/weights/best.pt --device 0 --save-conf --iou-thres 0.7  
val: data=./test.yaml, weights=['./yolov5/runs/train/piano_yolov5s_results/weights/best.pt'], bat ch_size=32, imgsz=640, conf_thres=0.001, iou_thres=0.7, max_det=300, task=test, device=0, workers =8, single_cls=False, augment=False, verbose=False, save_txt=False, save_hybrid=False, save_conf= True, save_json=False, project=yolov5\runs\val, name=exp, exist_ok=False, half=False, dnn=False fatal: cannot change to 'C:\Users\User\Desktop\fingers': No such file or directory Y0L0v5 2023-8-3 Python-3.10.11 torch-2.0.1 CUDA:0 (NVIDIA GeForce RTX 4080, 16376MiB)
```

The image below is the result of the YOLOv5 model's evalutaion metric.

```
YOLOv5s summary: 157 layers, 7012822 parameters, 0 gradients, 15.8 GFLOPs
test: Scanning C:\Users\User\Desktop\fingers detection\comparison_result_2023-08-14\YOLOv5\keybo
Class Images Instances P R mAP50 mAP50-95: 100%|
all 261 261 1 1 0.995 0.814
Speed: 0.2ms pre-process, 3.6ms inference, 2.1ms NMS per image at shape (32, 3, 640, 640)
```

Precision: 1

• Recall: 1

• mAP50: 0.995

The image below is the command to get YOLOv5 model's evalutaion metric in cmd.

```
(base) C:\Users\User\Desktop\fingers detection\comparison_result_2023-08-14\Y0L0v5>python ./yolov 5/val.py --task "test" --data ./test.yaml --weights ./yolov5/runs/train/piano_yolov5s_results/weights/best.pt --device 0 --save-conf --iou-thres 0.8 val: data=./test.yaml, weights=['./yolov5/runs/train/piano_yolov5s_results/weights/best.pt'], bat ch_size=32, imgsz=640, conf_thres=0.001, iou_thres=0.8, max_det=300, task=test, device=0, workers =8, single_cls=False, augment=False, verbose=False, save_txt=False, save_hybrid=False, save_conf= True, save_json=False, project=yolov5\runs\val, name=exp, exist_ok=False, half=False, dnn=False fatal: cannot change to 'C:\Users\User\Desktop\fingers': No such file or directory Y0L0v5 2023-8-3 Python-3.10.11 torch-2.0.1 CUDA:0 (NVIDIA GeForce RTX 4080, 16376MiB)
```

The image below is the result of the YOLOv5 model's evalutaion metric.

```
YOLOv5s summary: 157 layers, 7012822 parameters, 0 gradients, 15.8 GFLOPs
test: Scanning C:\Users\User\Desktop\fingers detection\comparison_result_2023-08-14\YOLOv5\keybo
Class Images Instances P R mAP50 mAP50-95: 100%|
all 261 261 1 1 0.995 0.811
Speed: 0.2ms pre-process, 3.7ms inference, 1.7ms NMS per image at shape (32, 3, 640, 640)
```

Precision: 1

• Recall: 1

• mAP50: 0.995

The image below is the command to get YOLOv5 model's evalutaion metric in cmd.

```
(base) C:\Users\User\Desktop\fingers detection\comparison_result_2023-08-14\V0L0v5>python ./yolov 5/val.py --task "test" --data ./test.yaml --weights ./yolov5/runs/train/piano_yolov5s_results/weights/best.pt --device 0 --save-conf --iou-thres 0.9 val: data=./test.yaml, weights=['./yolov5/runs/train/piano_yolov5s_results/weights/best.pt'], bat ch_size=32, imgsz=640, conf_thres=0.001, iou thres=0.9 max_det=300, task=test, device=0, workers =8, single_cls=False, augment=False, verbose=False, save_txt=False, save_hybrid=False, save_conf=True, save_json=False, project=yolov5\runs\val, name=exp, exist_ok=False, half=False, dnn=False fatal: cannot change to 'C:\Users\User\Desktop\fingers': No such file or directory Y0L0v5 2023-8-3 Python-3.10.11 torch-2.0.1 CUDA:0 (NVIDIA GeForce RTX 4080, 16376MiB)
```

The image below is the result of the YOLOv5 model's evalutaion metric.

```
YOLOV5s summary: 157 layers, 7012822 parameters, 0 gradients, 15.8 GFLOPs
test: Scanning C:\user\User\Desktop\fingers detection\comparison_result_2023-08-14\YOLOV5\keybo
Class Images Instances P R mAP50 mAP50-95: 100%|
all 261 261 0.995 1 0.995 0.813
Speed: 0.2ms pre-process, 3.9ms inference, 2.0ms NMS per image at shape (32, 3, 640, 640)
```

Precision: 0.995

• Recall: 1

mAP50: 0.995
mAP50-95: 0.813

The image below is the command to get YOLOv5 model's evalutaion metric in cmd.

```
(base) C:\Users\User\Desktop\fingers detection\comparison_result_2023-08-14\Y0L0v5>python ./yolov 5/val.py --task "test" --data ./test.yaml --weights ./yolov5/runs/train/piano_yolov5s_results/weights/best.pt --device 0 --save-conf --iou-thres 0.95 val: data=./test.yaml, weights=['./yolov5/runs/train/piano_yolov5s_results/weights/best.pt'], bat ch_size=32, imgsz=640, conf_thres=0.001, iou_thres=0.95, max_det=300, task=test, device=0, worker s=8, single_cls=False, augment=False, verbose=False, save_txt=False, save_hybrid=False, save_conf=True, save_json=False, project=yolov5\runs\val, name=exp, exist_ok=False, half=False, dnn=False fatal: cannot change to 'C:\Users\User\Desktop\fingers': No such file or directory Y0L0v5 2023-8-3 Python-3.10.11 torch-2.0.1 CUDA:0 (NVIDIA Geforce RTX 4080, 16376MiB)
```

The image below is the result of the YOLOv5 model's evalutaion metric.

```
YOLOv5s summary: 157 layers, 7012822 parameters, 0 gradients, 15.8 GFLOPs
test: Scanning C:\Users\User\Desktop\fingers detection\comparison_result_2023-08-14\Y0L0v5\keybo
Class Images Instances P mAP50 mAP50-95: 100%|
all 261 0.951 0.989 0.992 0.81
Speed: 0.2ms pre-process, 3.2ms inference, 1.0ms NMS per image at shape (32, 3, 640, 640)
```

• Precision: 0.951

• Recall: 0.989

• mAP50: 0.992

2. Evaluation Metric for Object Detection

- SSD-MobileNet

2. Evaluation Metric for Object Detection — SSD-MobileNet

- I coded to create a graph about PR curve using numpy, matplotlib, seaborn and scikit-learn.
- Used IoU threshold are 0.5, 0.6, 0.7, 0.8, 0.9 and 0.95.

import numpy as np
from sklearn.metrics import auc
import matplotlib.pyplot as plt
import seaborn as sns

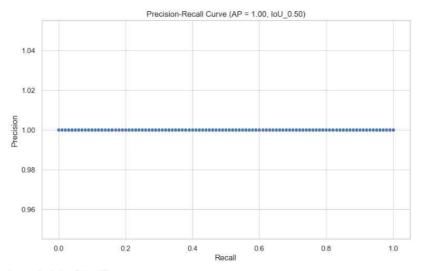
#output jou 0.50 # Calculate mAP by taking the average of AP values for all classes precision_values = ['1.00', '1 recall_values = [10.001, 10.011, 10.011, 10.021, 10.021, 10.021, 10.031, 10.031, precision_values = [float(value) for value in precision_values] recall_values = [float(value) for value in recall_values] precision = np.array(precision_values, dtype=float) recall = np.array(recall_values, dtype=float) # Calculate AP using the area under the precision-recall curve ap = auc(recall, precision) # Calculate AR by averaging recall values ar = np.mean(recall) # Calculate mAP by taking the average of AP values for all classes # Set seaborn style sns.set(style="whitegrid") # Pint Pracision-Recall curve plt.figure(figsize=(10, 6)) sns.lineplot(x=recall, y=precision, marker='o', color='b') plt.xlahel('Recall') plt.ylabel('Precision') plt.title('Precision-Recall Curve (AP = {:.2f}, IoU_0.50)', format(ap)) plt.grid(True) plt.show() print("Average Precision (AP):", ap) print("Average Recall (AR):", ar) print("Mean Average Precision (mAP):", mAP)

2. Evaluation Metric Summary - SSD-MobileNet

```
Calculating mAP at 0.50 foll threshold...
100.00% = keyboard AP
MAP = 100.00%
Calculating mAP at 0.55 loU threshold...
100.00% = keyboard AP
MAP = 100.00%
Calculating mAP at 0.60 loU threshold...
100.00% = keyboard AP
MAP = 100.00%
Calculating mAP at 0.65 loU threshold...
99.43% = keyboard AP
MAP = 99.43\%
Calculating mAP at 0.70 loU threshold...
99.01% = keyboard AP
mAP = 99.01\%
Calculating mAP at 0.75 loU threshold...
96.91% = keyboard AP
mAP = 96.91\%
```

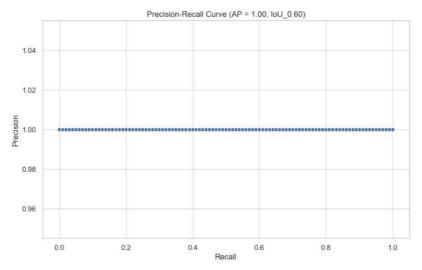
```
Calculating mAP at 0.80 loU threshold...
83.74% = keyboard AP
MAP = 83.74\%
Calculating mAP at 0.85 loU threshold...
53.59% = keyboard AP
mAP = 53.59\%
Calculating mAP at 0.90 loU threshold...
16.38% = kevboard AP
mAP = 16.38\%
Calculating mAP at 0.95 loU threshold...
0.48% = keyboard AP
mAP = 0.48\%
***MAP Results***
                Average mAP @ 0.5:0.95
Class
                        74.95%
kevboard
Overall
                74.95%
```

• The image below is the result of the SSD-MobileNet model's evalutaion metric.



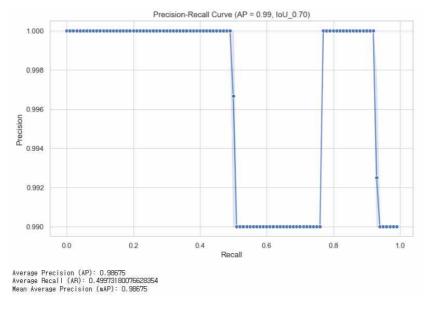
Average Precision (AP): 1.0 Average Recall (AR): 0.5019157088122606 Mean Average Precision (mAP): 1.0

• The image below is the result of the SSD-MobileNet model's evalutaion metric.

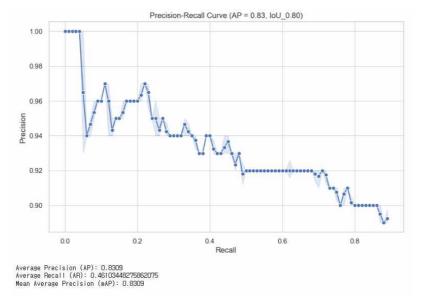


Average Precision (AP): 1.0 Average Recall (AR): 0.5019157088122606 Mean Average Precision (mAP): 1.0

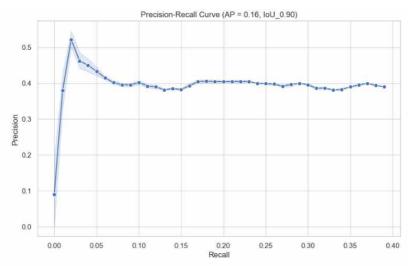
• The image below is the result of the SSD-MobileNet model's evalutaion metric.



• The image below is the result of the SSD-MobileNet model's evalutaion metric.

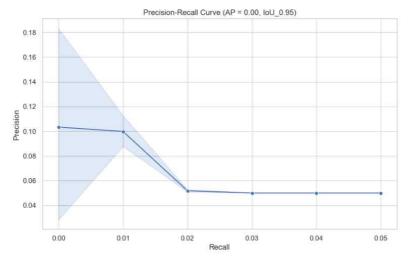


• The image below is the result of the SSD-MobileNet model's evalutaion metric.



Average Precision (AP): 0.1557 Average Recall (AR): 0.19831417624521072 Mean Average Precision (mAP): 0.1557

• The image below is the result of the SSD-MobileNet model's evalutaion metric.



Average Precision (AP): 0.0043 Average Recall (AR): 0.026245210727969343 Mean Average Precision (mAP): 0.0043

2. Evaluation Metric for Object Detection - SSD-MobileNet

• Files about evalution metric of SSD-MobileNet are attached in this folder with the name 'mAP(SSD).

How to understanding those



lcp /content/images/test/* /content/mAP/input/images-optional # Copy images and xml files
!mv /content/mAP/input/images-optional/*.xml /content/mAP/input/ground-truth/ # Move xml files to the appropriate folder

outputs: each output files about IoU threshold values

2. Evaluation Metric for Object Detection - SSD-MobileNet

- · First, We need to enter the outputs directory.
- · Second, open the output txt file in the output iou folder.
- This image below is example about IoU threshold 0.5.

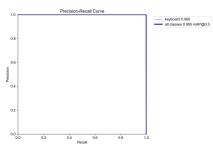
```
# AP and precision/recall per class
100.00% = keyboard AP
  Precision: ['1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1
'1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00'
 '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00',
 '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00'
 '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00'
 '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00'
 '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00'
'1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00'
 '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00', '1.00'
  Recall :['0.00', '0.01', '0.01', '0.02', '0.02', '0.02', '0.0
 '0.11', '0.11', '0.12', '0.12', '0.13', '0.13', '0.13', '0.14'
 '0.22', '0.23', '0.23', '0.23', '0.24', '0.24', '0.25', '0.25'
 '0.33', '0.34', '0.34', '0.34', '0.35', '0.35', '0.36', '0.36'
 '0.44', '0.45', '0.45', '0.46', '0.46', '0.46', '0.47', '0.47'
 '0.56', '0.56', '0.56', '0.57', '0.57', '0.57', '0.58', '0.58'
 '0.67', '0.67', '0.67', '0.68', '0.68', '0.69', '0.69', '0.69'
 '0.78', '0.78', '0.79', '0.79', '0.79', '0.80', '0.80', '0.80'
 '0.89', '0.89', '0.90', '0.90', '0.90', '0.91', '0.91', '0.92'
 '1.00'1
# mAP of all classes
 mAP = 100.00\%
# Number of ground-truth objects per class
keyboard: 261
# Number of detected objects per class
keyboard: 261 (tp:261, fp:0)
```

- Precision List: precision values for test dataset in './input/images-optional'
- Recall List: recall values for test dataset in ',/input/images—optional'
- mAP: mAP value found at each IoU value.
- * If you want to understand something additionally, please refer to the 'calculate_map_cartucho.py' and 'main.py' code attached.

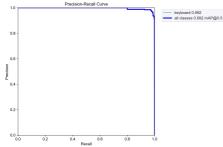
3. Comparison the results (PR-Curve)

3. Comparison the results (PR-Curve)

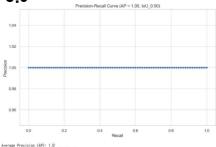
YOLOv5 ~ IoU: 0.5



• YOLOv5 ~ IoU: 0.95

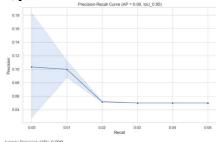


• SSD ~ IoU: 0.5



Average Recail (AR): 0.5019157088122506 Mean Average Precision (mAP): 1.0

• SSD ~ IoU: 0.95



Average Recall (AR): 0.025245210727969

3. Analysis of evalutaion metric results

YOLOv5 & SSD-MobileNet

- · As explained earlier, I collected 2,605 images from divided in 5 technique related videos. Then I trained the model using these.
- Although I split dataset for train, test, and val respectively, For example, elements such as the background are too similar, so I think overfitting occurs. Also, Precision and Recall values are too high.

YOLOv5 vs SSD-MobileNet

I think the model trained YOLOv5 is better than another.