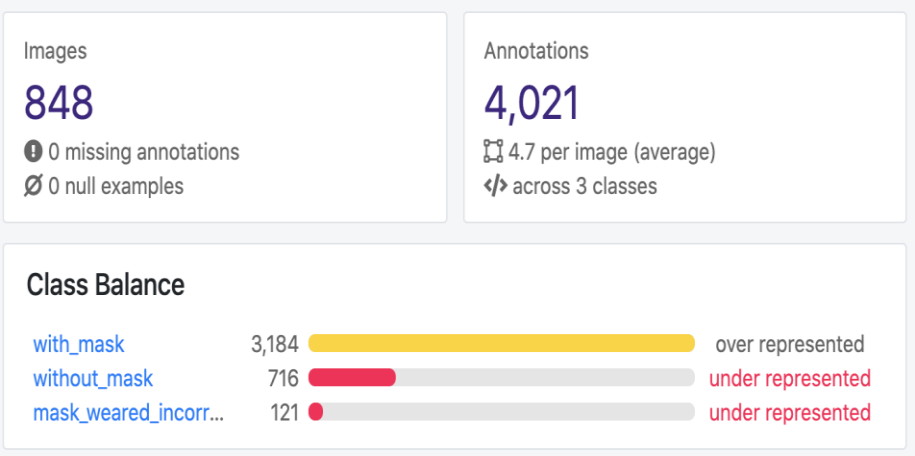
Facemask Detection

Conquer Covid-19

1. Almost a year later, the world still has about 70 million known cases and over 600 thousand new cases a day!
2. Although the vaccine is on the way, wearing mask and social distancing are the most effective ways to stop the spread in the meantime.
3. Social distancing is difficult to control, but enforcing face masks to be worn can be assisted through DL technologies to quickly identify if individuals are wearing face masks. And if so, are the masks being worn properly?

Original Dataset

1. https://www.kaggle.com/andrewmvd/face-mask-detection
2. 853 images, each has one or multiple faces, wearing masks, no masks or wearing masks incorrectly
3. Each image is pre-labeled to show the x-y coordinates of the bounding box of each face and its classification - no mask, mask worn correctly, or mask worn incorrectly



Original Dataset examples


```
<annotation>
    <folder>images</folder>
    <filename>maksssksksss2.png</filename>
    <size>
        <width>400</width>
        <height>290</height>
        <depth>3</depth>
    </size>
    <segmented>0</segmented>
    <object>
        <name>with_mask</name>
        <pose>Unspecified</pose>
        <truncated>0</truncated>
        <occluded>0</occluded>
        <difficult>0</difficult>
        <bndbox>
            <xmin>68</xmin>
            <ymin>42</ymin>
            <xmax>105</xmax>
            <ymax>69</ymax>
        </bndbox>
    </object>
    <object>
        <name>with_mask</name>
        <pose>Unspecified</pose>
        <truncated>0</truncated>
        <occluded>0</occluded>
        <difficult>0</difficult>
        <bndbox>
            <xmin>154</xmin>
            <ymin>47</ymin>
            <xmax>178</xmax>
            <ymax>74</ymax>
        </bndbox>
    </object>
    <object>
        <name>with_mask</name>
        <pose>Unspecified</pose>
        <truncated>0</truncated>
        <occluded>0</occluded>
        <difficult>0</difficult>
        <bndbox>
            <xmin>238</xmin>
            <ymin>34</ymin>
            <xmax>262</xmax>
            <ymax>69</ymax>
        </bndbox>
    </object>
    <object>
        <name>mask_worn_incorrect</name>
```

Original Dataset examples

```
<annotation>
    <folder>images</folder>
    <filename>maksssksksss0.png</filename>
    <size>
        <width>512</width>
        <height>366</height>
        <depth>3</depth>
    </size>
    <segmented>0</segmented>
    <object>
        <name>without_mask</name>
        <pose>Unspecified</pose>
        <truncated>0</truncated>
        <occluded>0</occluded>
        <difficult>0</difficult>
        <bndbox>
            <xmin>79</xmin>
            <ymin>105</ymin>
            <xmax>109</xmax>
            <ymax>142</ymax>
        </bndbox>
    </object>
    <object>
        <name>with_mask</name>
        <pose>Unspecified</pose>
        <truncated>0</truncated>
        <occluded>0</occluded>
        <difficult>0</difficult>
        <bndbox>
            <xmin>185</xmin>
            <ymin>100</ymin>
            <xmax>226</xmax>
            <ymax>144</ymax>
        </bndbox>
    </object>
    <object>
        <name>without_mask</name>
        <pose>Unspecified</pose>
        <truncated>0</truncated>
        <occluded>0</occluded>
        <difficult>0</difficult>
        <bndbox>
            <xmin>325</xmin>
            <ymin>90</ymin>
            <xmax>360</xmax>
            <ymax>141</ymax>
        </bndbox>
    </object>
</annotation>
```

Dataset for Pre-Trained YOLO and Transfer Learning - Mobilenet

Class 0: No mask incorrectly

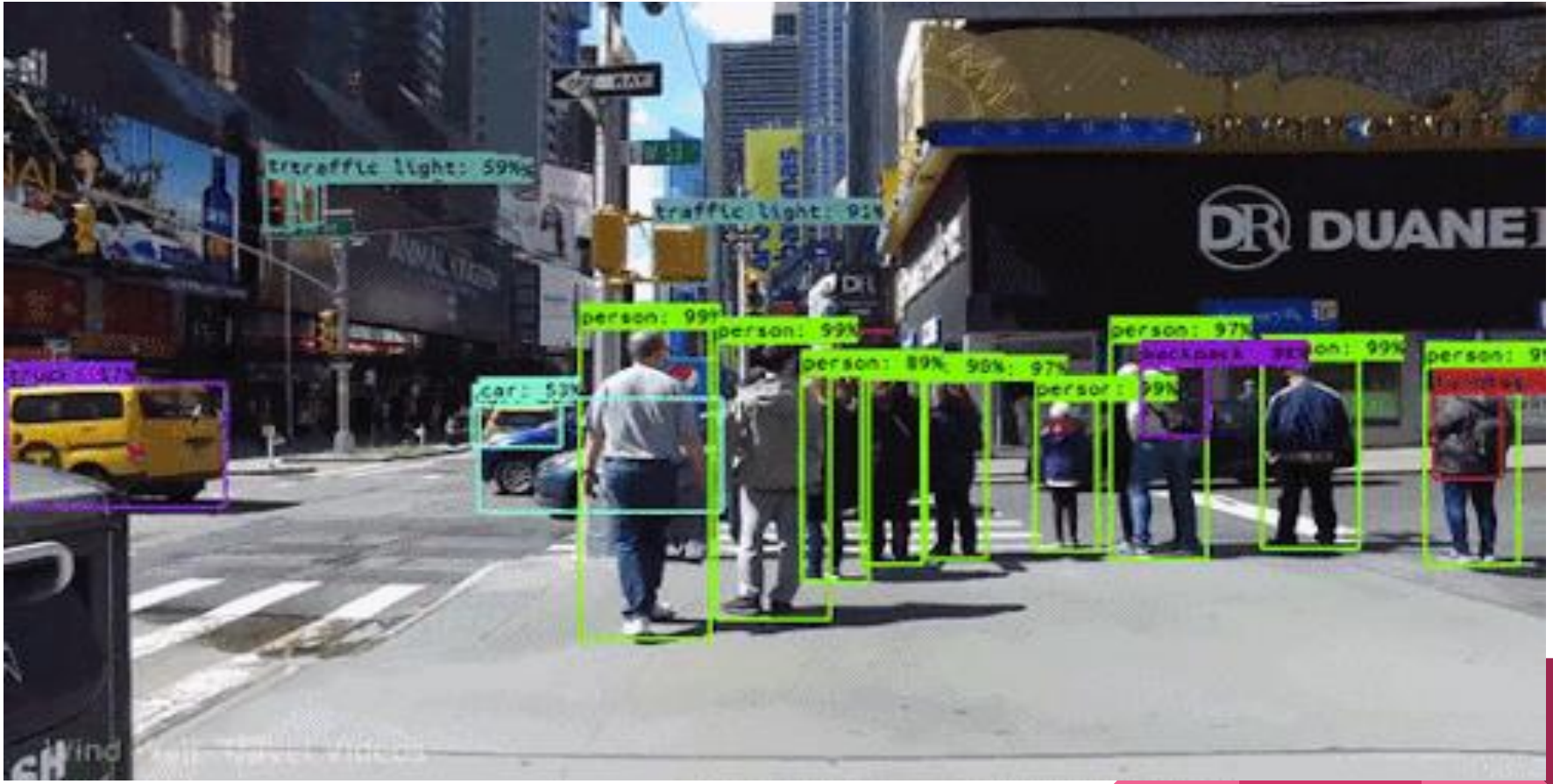
Class 1: Mask worn correctly

Class 2: Mask worn

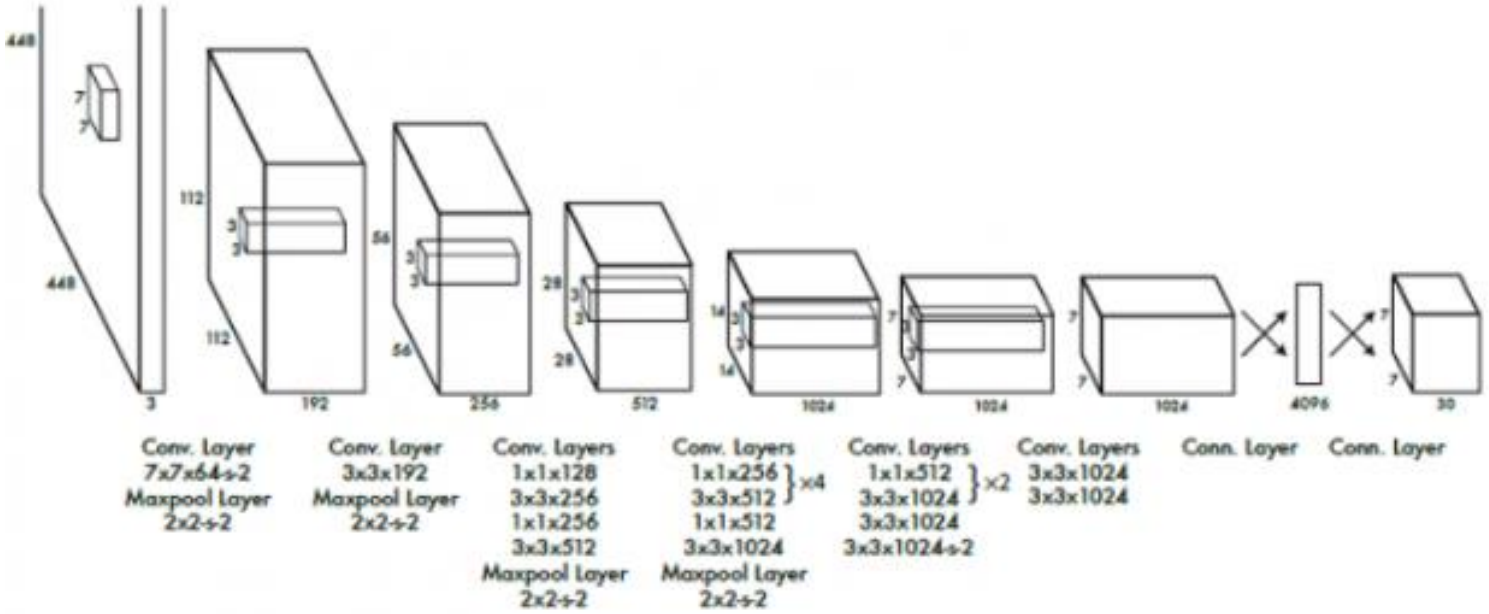
1. https://drive.google.com/drive/folders/1XDte2DL2Mf_hw4NsmGst7QtYoU7sMBVG
    a. Real life images
    b. 1915 images of faces; 1919 images of faces with masks
2. https://github.com/cabani/MaskedFace-Net
    a. 67,193 images with correctly masked faces; 66,900 images with incorrectly masked faces. We only use images with incorrectly masked faces
    b. Mask are photoshopped.

Yolo (You Only Look Once)

1. https://arxiv.org/abs/1506.02640 2016 seminar paper by Joseph Redmon et al.
2. Real time objection detection algorithm
3. Understand what and where are in the image
4. High accuracy and real time
5. A single CNN simultaneously predicts multiple bounding boxes and class probabilities for those boxes

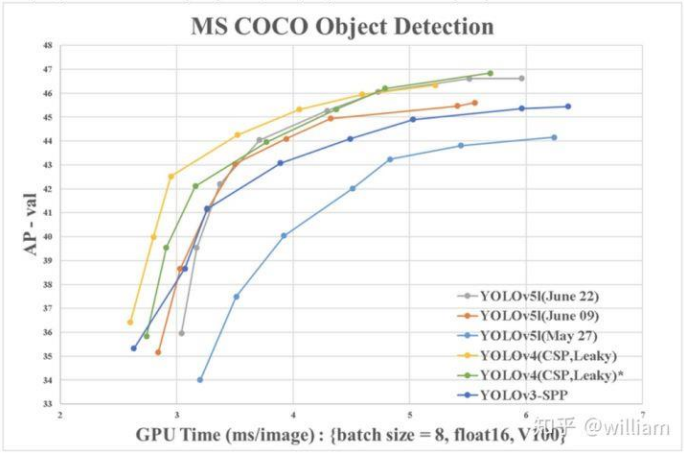

traffic light: 59%
traffic light: 9
person: 99%
person: 99%
person: 89%
98%: 97%
person: 97%
person: 99%
car: 53%
DUANE

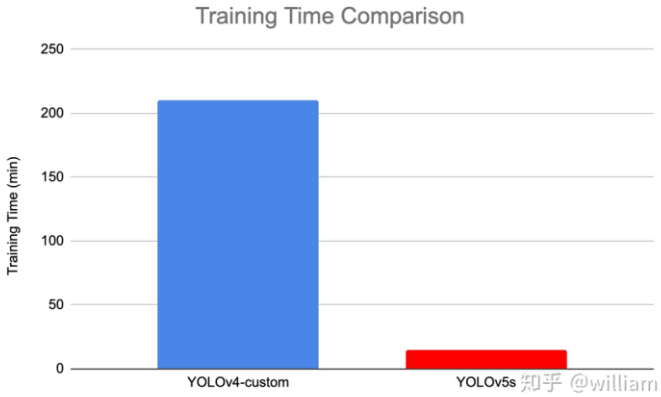
Yolo (You Only Look Once)

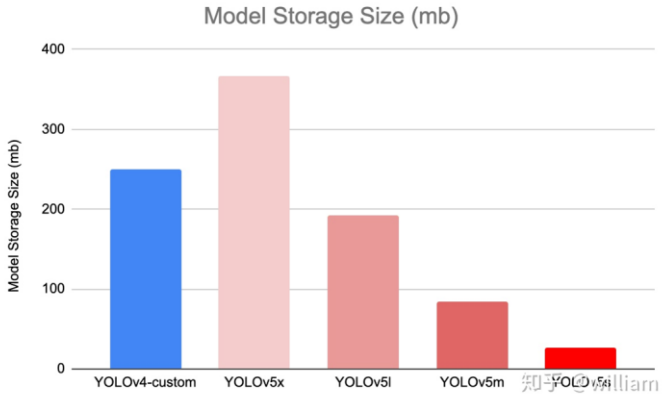


Yolo v5

1. A bit worse precision compared to Yolo v4
2. Much less training time
3. Less model storage size







Method 1 : Yolo + mobilenet

1. Use Yolo to detect all the faces in the image
2. Cropped all the detected faces
3. Train a CNN model that classifies faces into 3 categories using dataset 1
    a. Class 0 no mask
    b. Class 1 masked properly
    c. Class 2 masked incorrectly
4. Send all the cropped faces to classification model

Method 1 : Training and results

1. Use mobilenet as the base model and add the final layer to classifies the extracted features from mobilenet
2. Training dataset size 4,428. Testing dataset size 1,200
3. The testing dataset accuracy can hit 99% after 15 epochs.

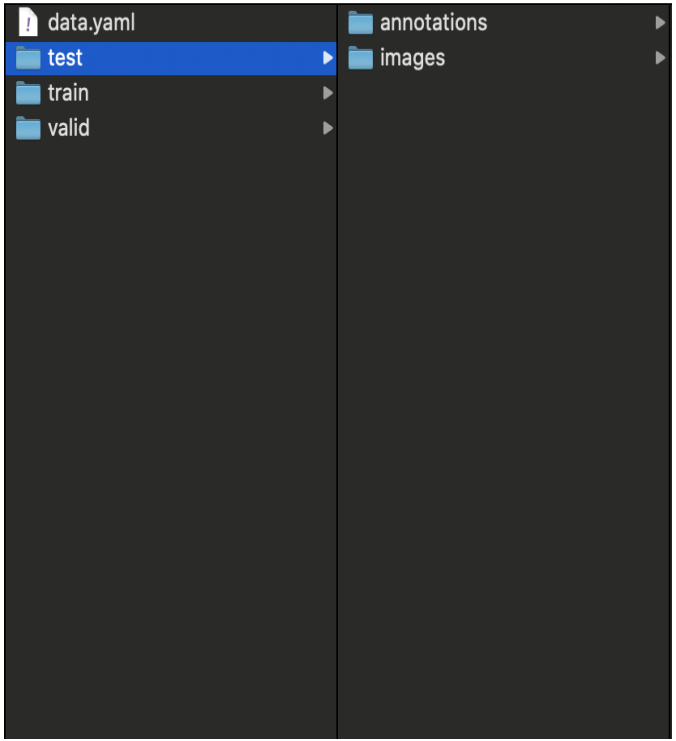
Method 2 : Data Preprocessing

```xml
<annotation>
    <folder>images</folder>
    <filename>maksssksksss0.png</filename>
    <size>
        <width>512</width>
        <height>366</height>
        <depth>3</depth>
    </size>
    <segmented>0</segmented>
    <object>
        <name>without_mask</name>
        <pose>Unspecified</pose>
        <truncated>0</truncated>
        <occluded>0</occluded>
        <difficult>0</difficult>
        <bndbox>
            <xmin>79</xmin>
            <ymin>105</ymin>
            <xmax>109</xmax>
            <ymax>142</ymax>
        </bndbox>
    </object>
    <object>
        <name>with_mask</name>
        <pose>Unspecified</pose>
        <truncated>0</truncated>
        <occluded>0</occluded>
        <difficult>0</difficult>
        <bndbox>
            <xmin>185</xmin>
            <ymin>100</ymin>
            <xmax>226</xmax>
            <ymax>144</ymax>
        </bndbox>
    </object>
    <object>
        <name>without_mask</name>
        <pose>Unspecified</pose>
        <truncated>0</truncated>
        <occluded>0</occluded>
        <difficult>0</difficult>
        <bndbox>
            <xmin>325</xmin>
            <ymin>90</ymin>
            <xmax>360</xmax>
            <ymax>141</ymax>
        </bndbox>
    </object>
</annotation>
```

maksssksksss0.txt

```
0 0.18359375 0.337431693989071 0.05859375 0.10109289617486339
2 0.4013671875 0.3333333333333333 0.080078125 0.12021857923497267
0 0.6689453125 0.3155737704918033 0.068359375 0.13934426229508196
```

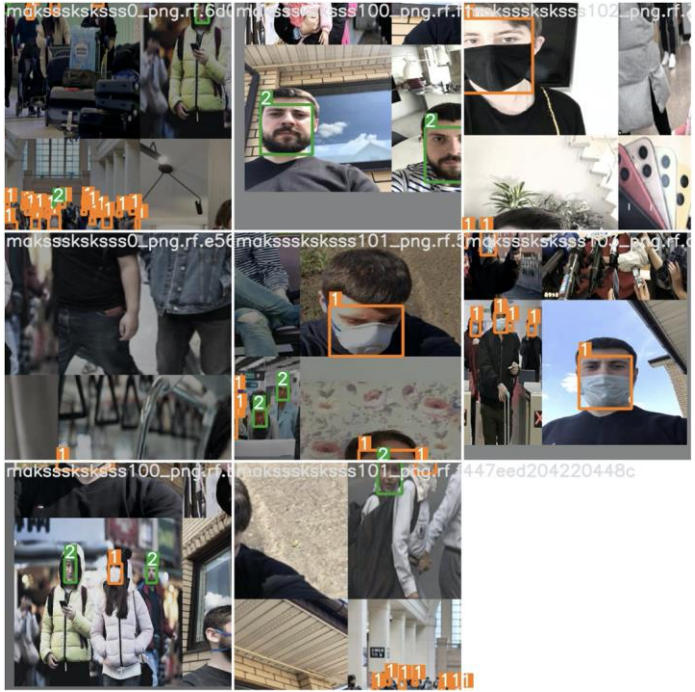
Method 2 : Data Preprocessing

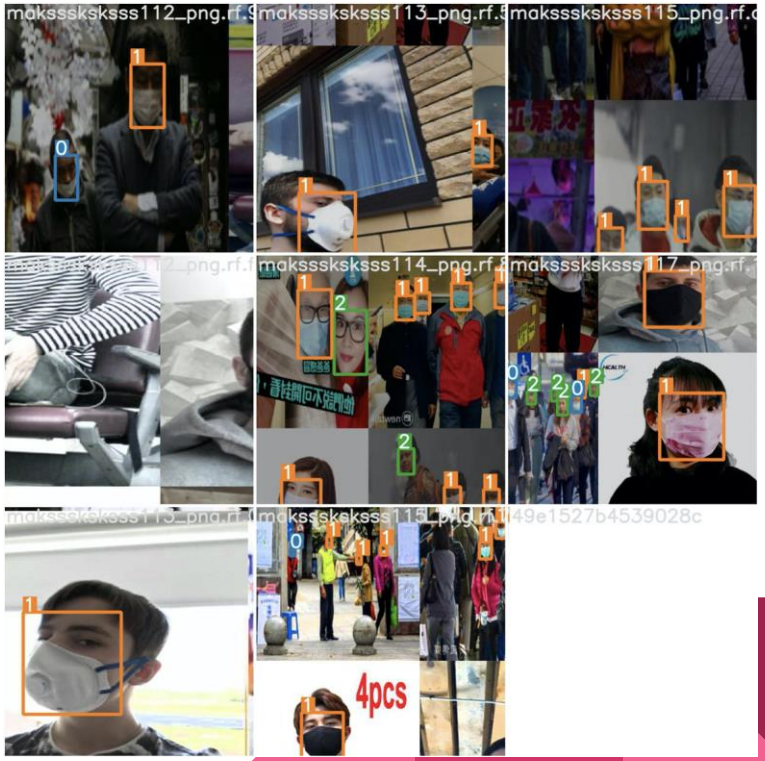



```
train: train/images
val: valid/images

nc: 3
names: ['without_mask', 'mask_weared_incorrect', 'with_mask']
```

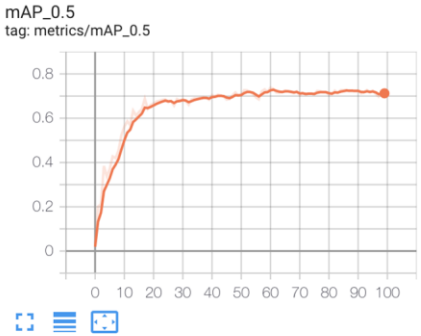
Method 2 : Training and results

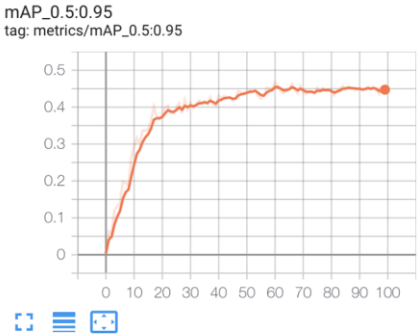


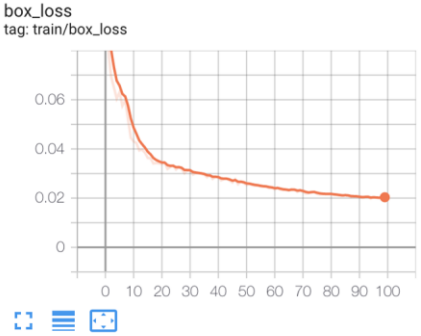


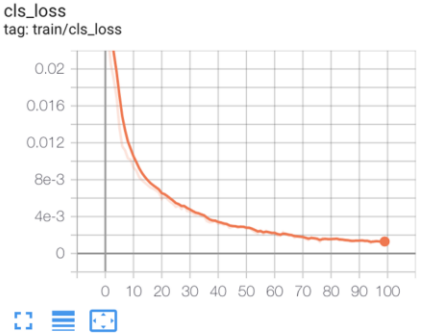
Method 2 : Training and results

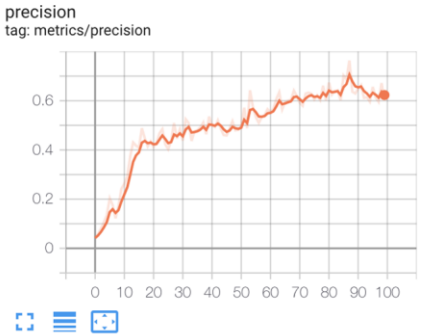
1. Use YOLO v5 detect different classes directly

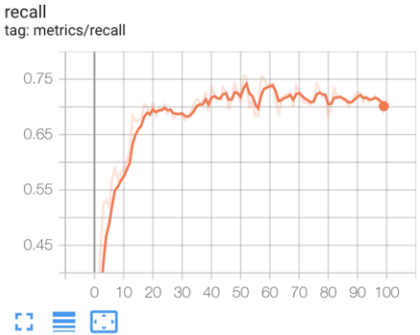


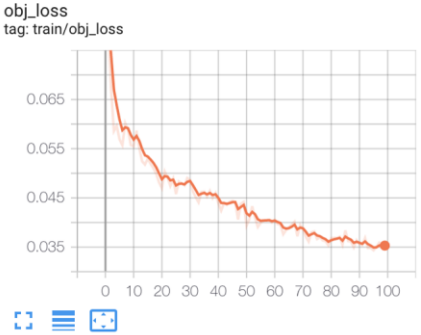




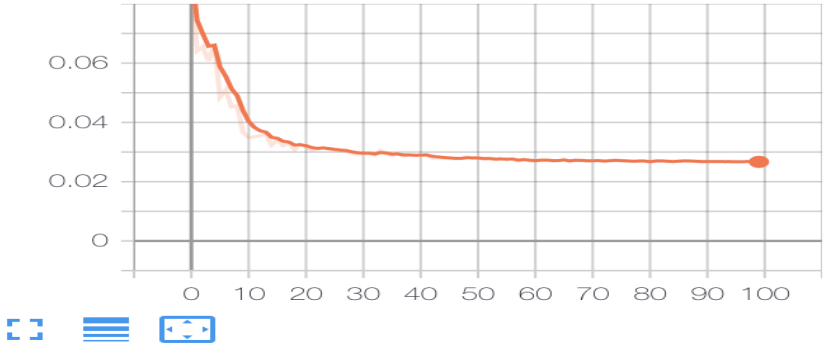




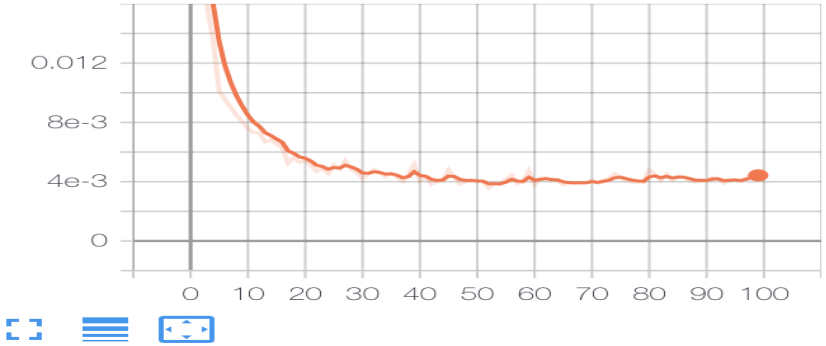




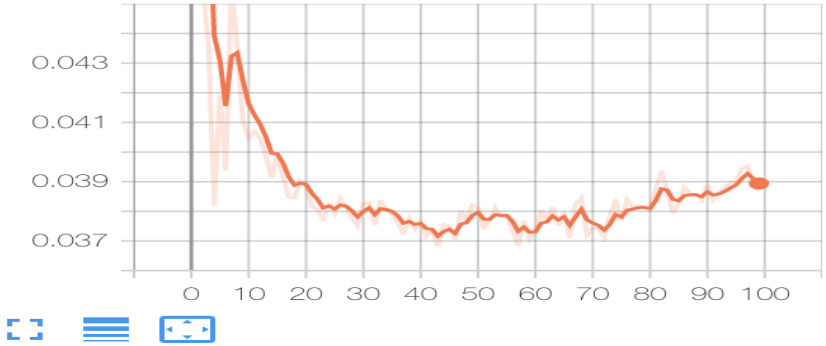
box_loss
tag: val/box_loss



cls_loss
tag: val/cls_loss



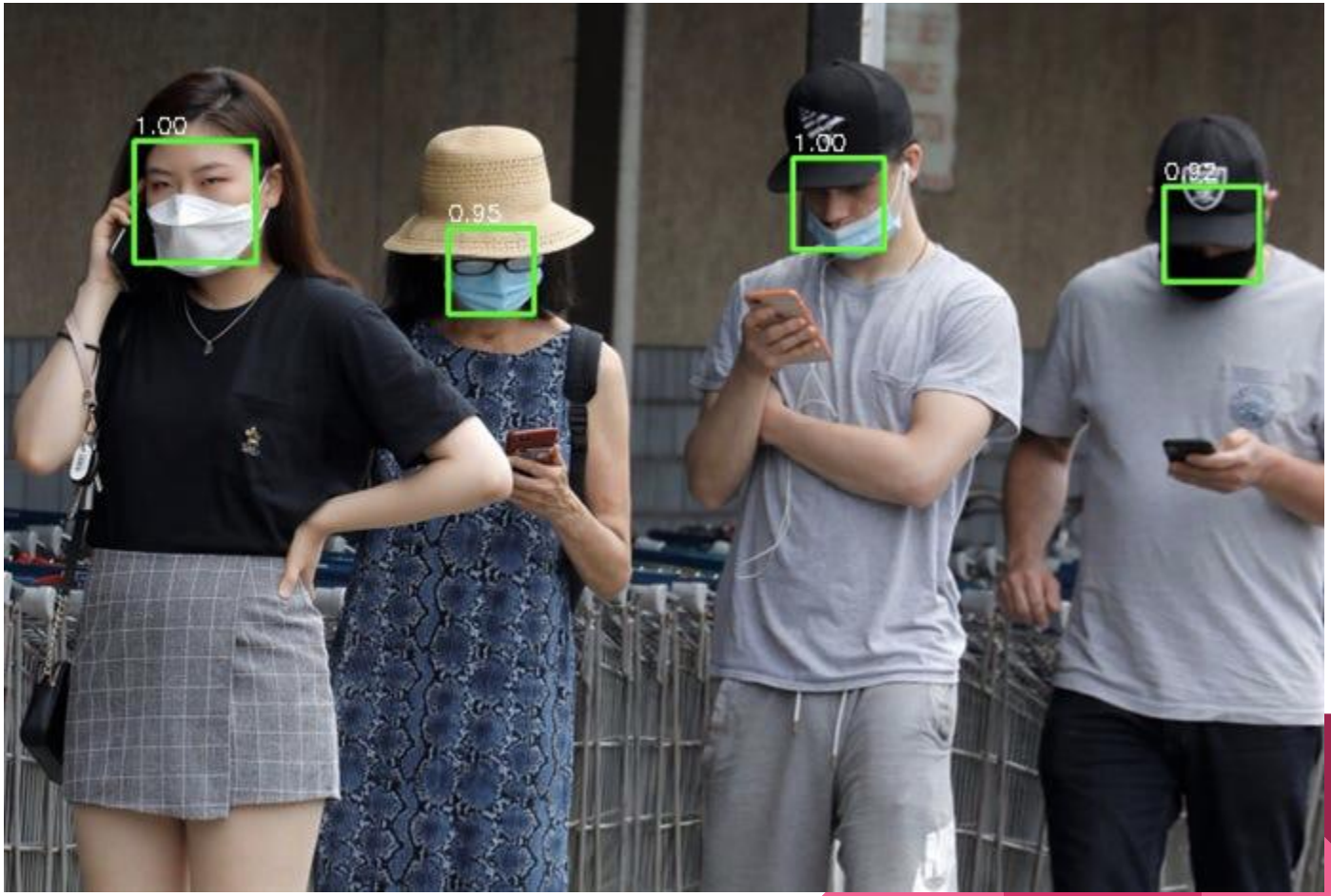
obj_loss
tag: val/obj_loss



Compare Method 1 and 2

	Number of picked up faces out of 9 images	Accuracy	Number of misclassifications: with_mask	Number of misclassifications: without_mask	Number of misclassifications: mask_worn_incorrectly
Method 1	57	50/57 ≈ 0.877	2	1	4
Method 2	86	81/86 ≈ 0.94	1	0	4

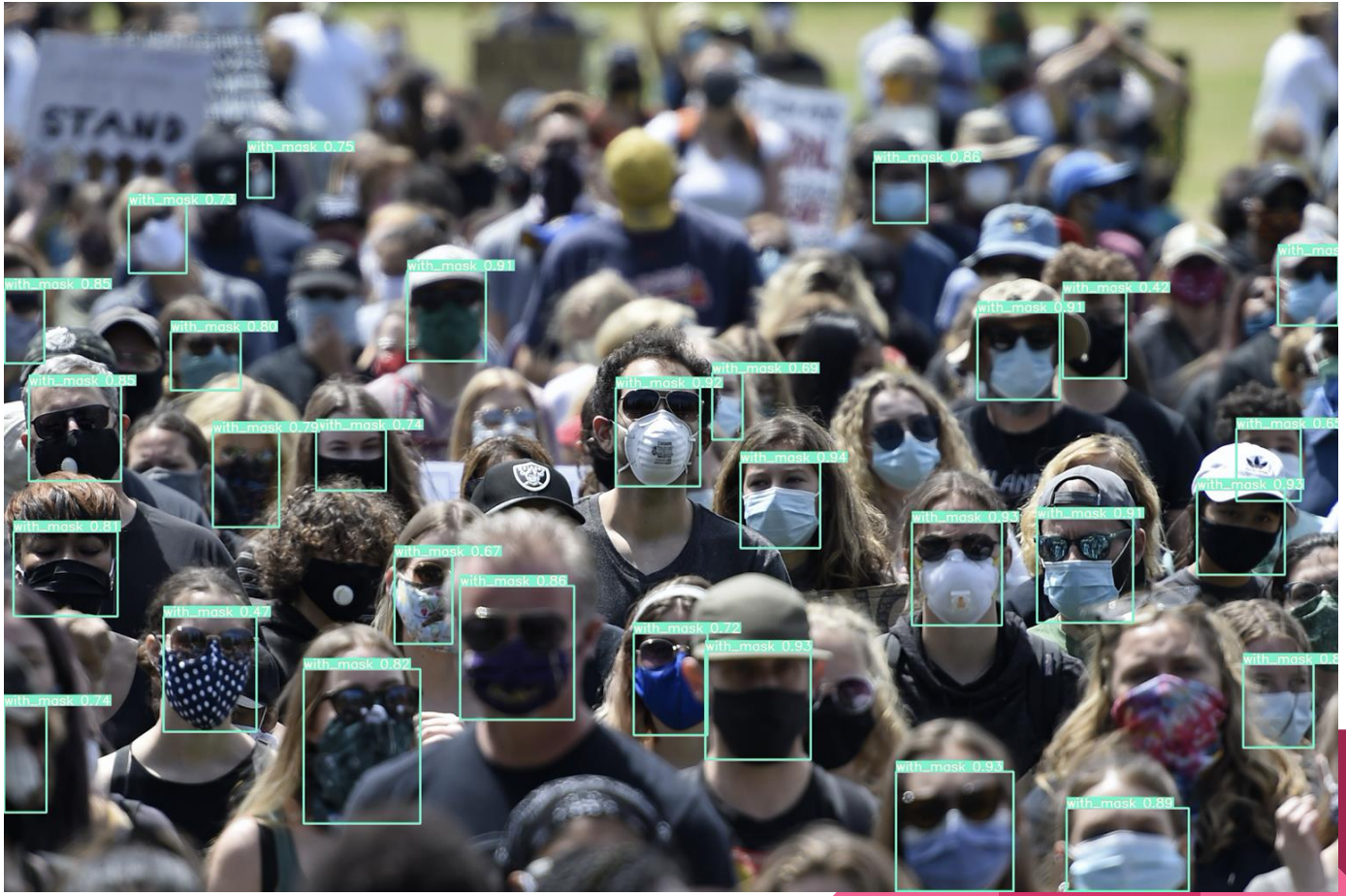
Method 1



Method 2

Method 1

Method 2


STAND
with_mask 0.75
with_mask 0.73
with_mask 0.86
with_mask 0.91
with_mask 0.85
with_mask 0.42
with_mask 0.91
with_mask 0.80
with_mask 0.69
with_mask 0.85
with_mask 0.74
with_mask 0.94
with_mask 0.93
with_mask 0.81
with_mask 0.91
with_mask 0.67
with_mask 0.86
with_mask 0.47
with_mask 0.72
with_mask 0.93
with_mask 0.82
with_mask 0.74
with_mask 0.93
with_mask 0.89

Future steps

1. Gather more data for classes: “without_mask” and “mask_worn_incorrectly” to balance dataset
2. Label the data
3. Retrain the model