



Mask Up

Face mask detection through deep learning



Intro & **Problem Statement**

Ongoing Pandemic

- Control of covid-19 pandemic
- Wearing mask is still the most effective way to inhibit virus transmission

Leverage tech to check for mask wearing compliance

- Use deep learning to allow computers to “see” the world around us
- Manpower could be deployed to help healthcare workers instead



How?

01.

Select Model

1. Single stage object detection model e.g. YOLO (you only look once)
2. Fast inference time, and high accuracy

03.

Training Model

1. Train the model so that it has good IoU score

02.

Collection of Dataset

1. Supervised problem
2. Sufficiently large dataset is required for training

04.

Deployment

1. Deploy trained model on the edge (IoT devices)

Data Collection and Cleaning



Data Collection

- Data available on Kaggle
- 848 images and annotation files



Duplicated images

- Using difPy, 5 duplicated images detected and removed (comparing MSE between 2 images)
- Important to remove duplicated images as they can cause data leakage



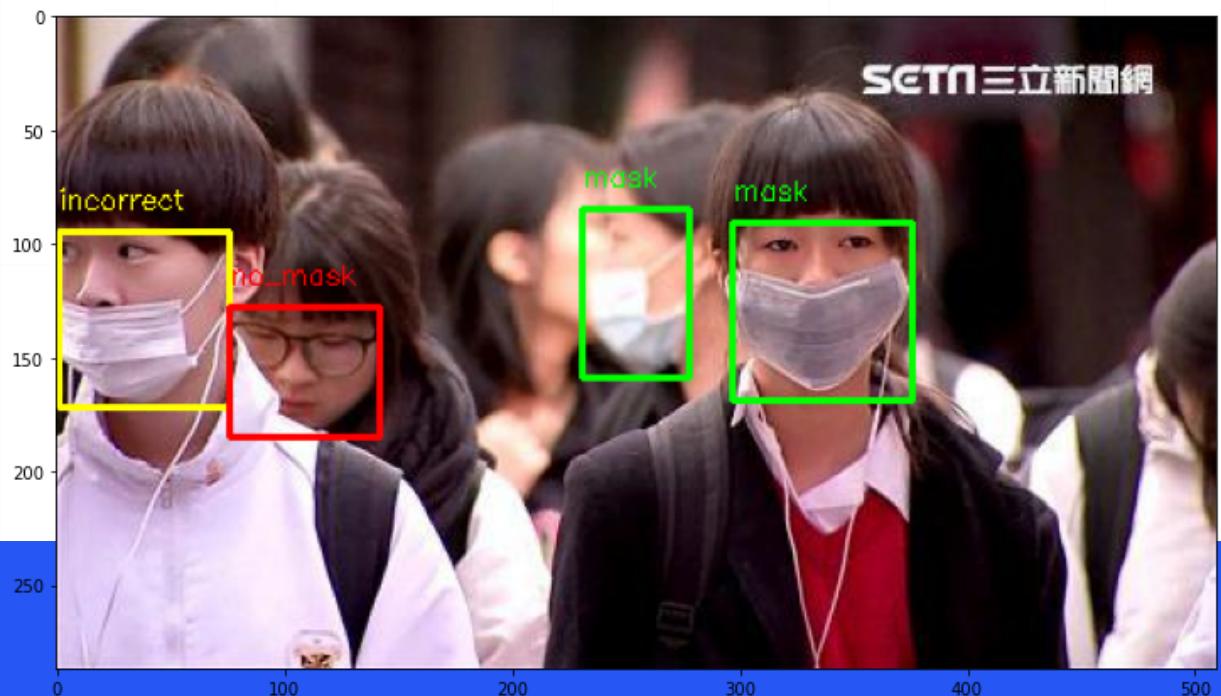
Change class label

- without_mask > no_mask
- with_mask > mask
- mask_weared_incorrect > incorrect

EDA

Check the quality of annotation

- Print out images and annotation information
- Check the bbox and class



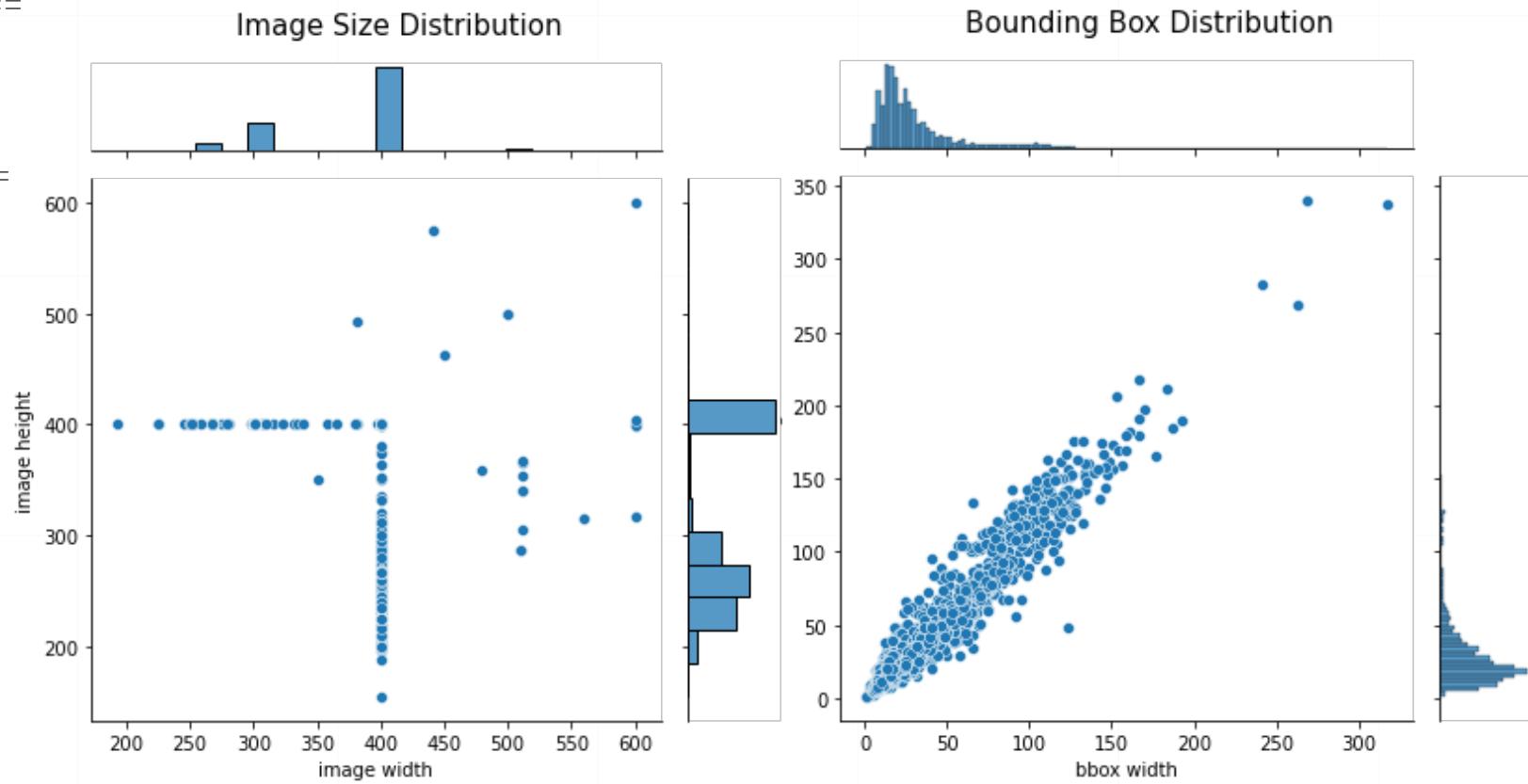
EDA

Check image size distribution

- ===== Bounding boxes statistics =====
- median width : 22.0
- median height: 24.0
- ===== Image statistics =====
- median width : 280.5
- median height: 280.5

Class distribution

- mask 0.79
- no_mask 0.18
- incorrect 0.03





Preprocessing - Roboflow image processing tool

Train, validation, split

- 70%, 20%, 10% split (593, 170, 85 images respectively)
- Class imbalance, hence split need to be stratify
- Roboflow does the split quite effortlessly ensuring that each class is represented in the correct proportion.



Image resizing

- Resized to square ratio, selected 640x640 and 320x320
- To maintained the aspect ration, padding is required. Black pixels used.





Preprocessing - Roboflow image processing tool

Image augmentation

- (1) horizontal flip, (2) $+10^\circ$ rotation, and (3) applied mosaic
- Increase the training data from 593 -> 1,779 (~ 3 times)
- Note that validation and test set remained at 170 and 85 images. Ground truth images, they should not be augmented.



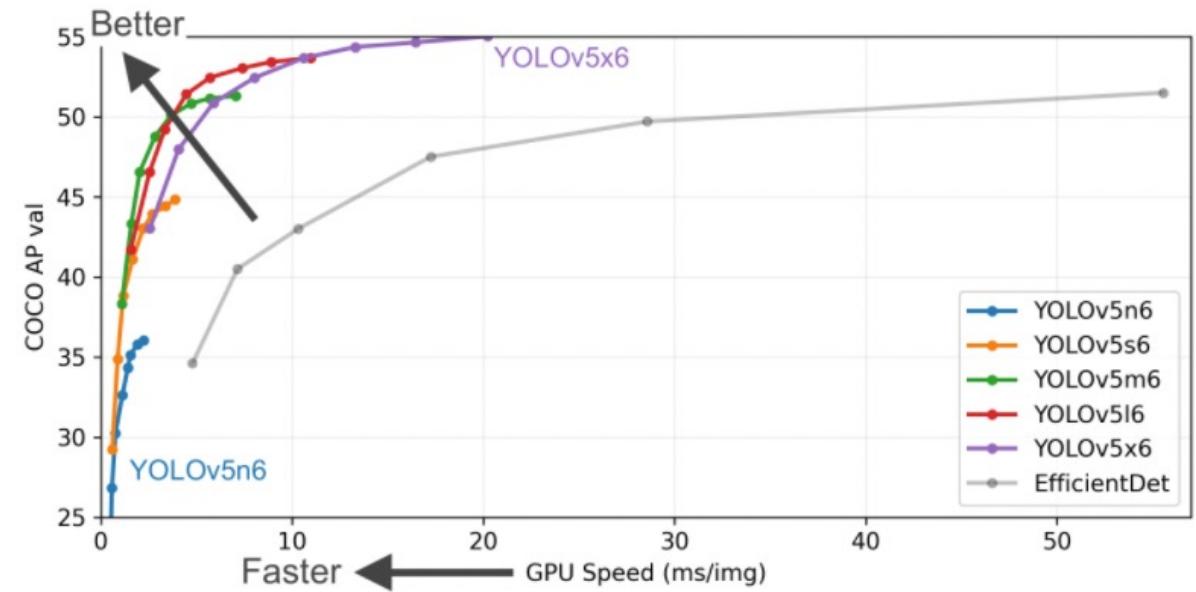
Modelling – YOLOv5

4 different runs

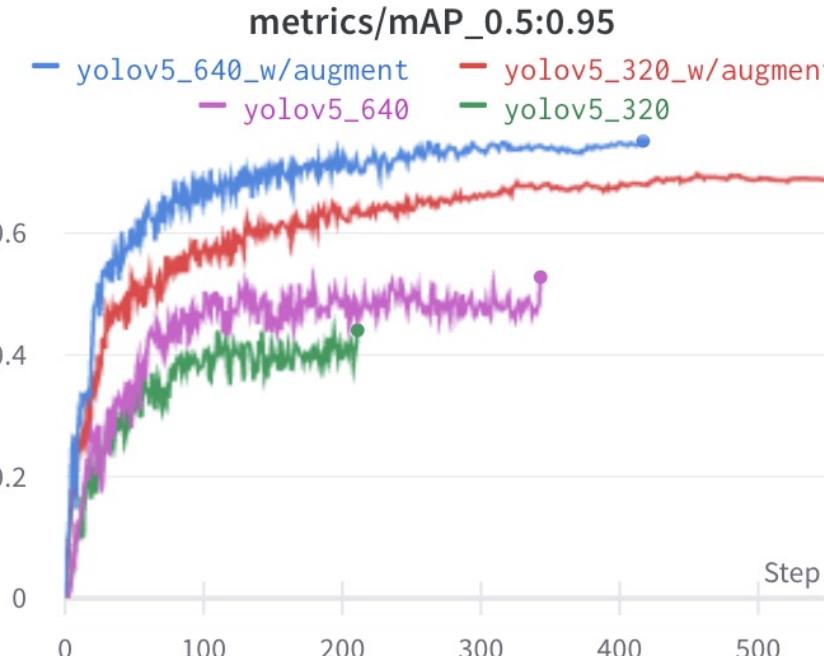
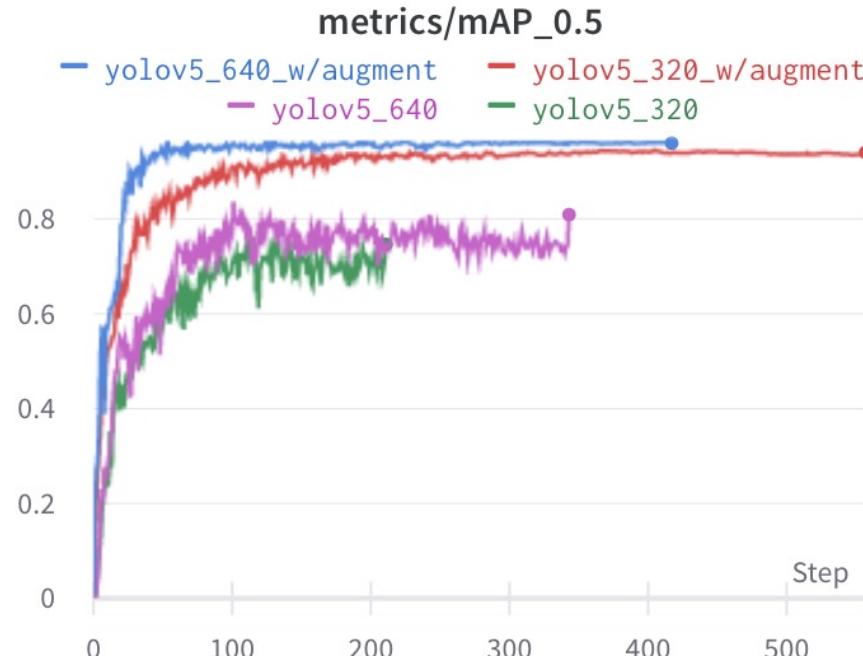
1. input image size 320, with no augmentations
2. input image size 320, with augmentations
3. input image size 640, with no augmentations
4. input image size 640, with augmentations

Transfer learning

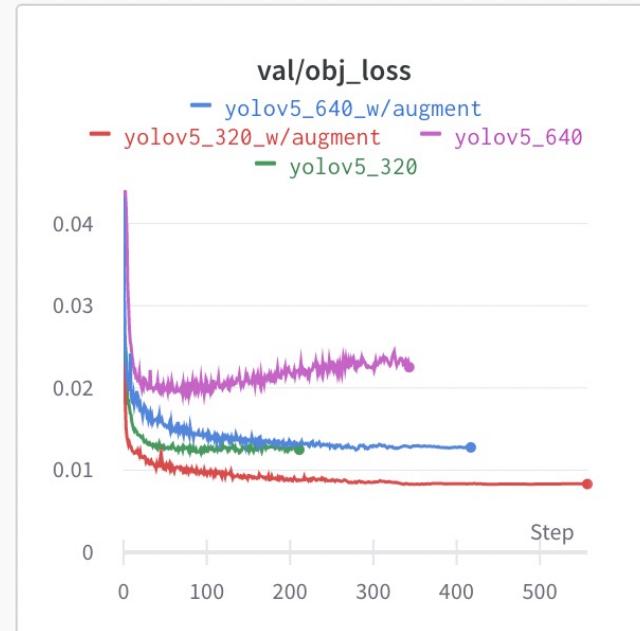
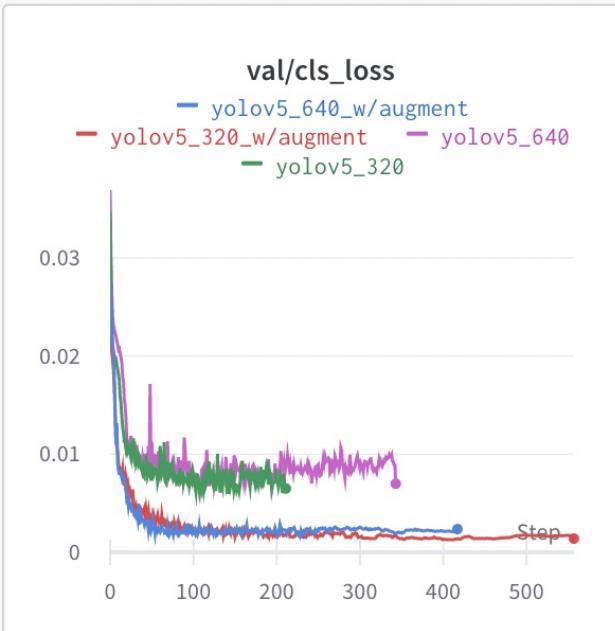
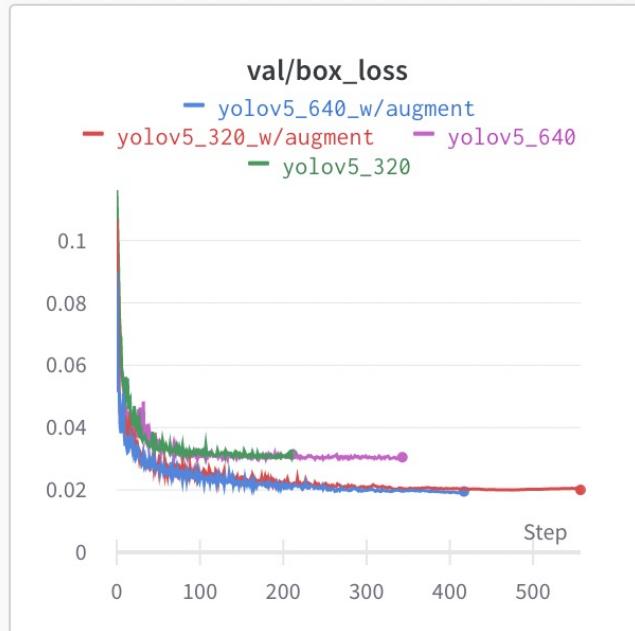
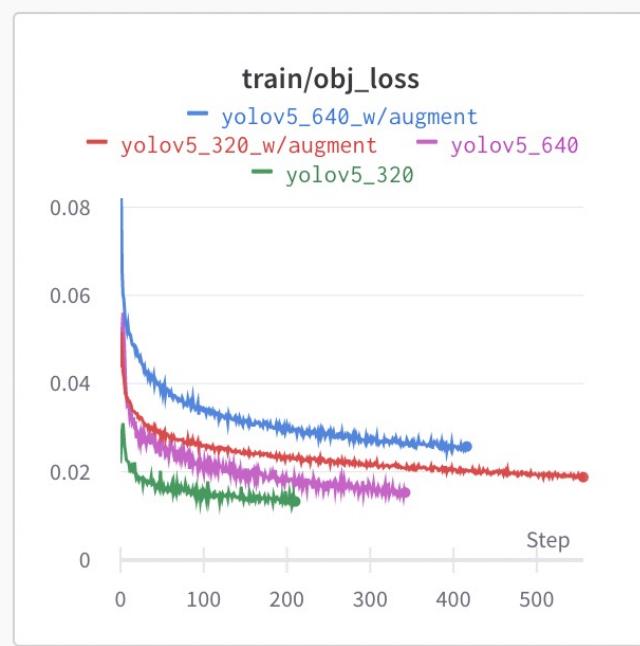
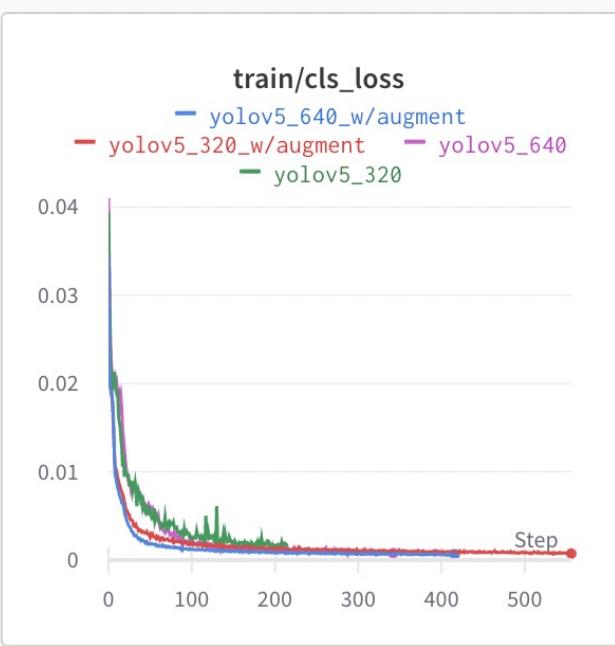
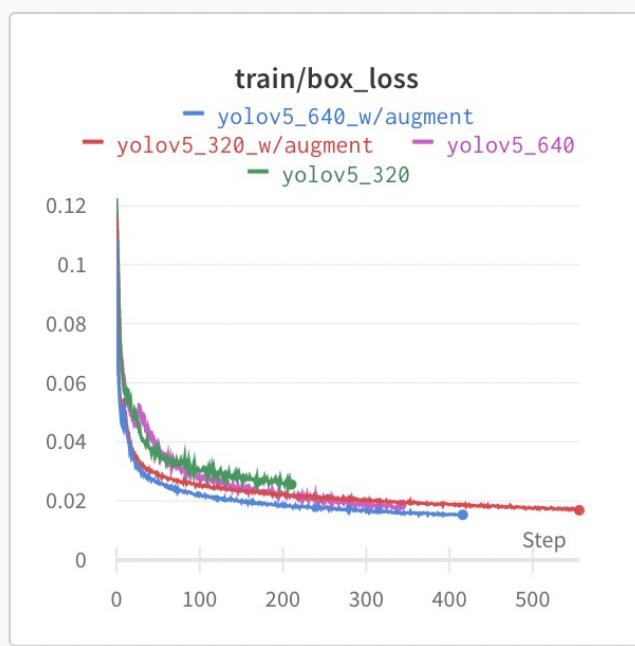
1. Pre-trained weights yolovs
2. "S" category is a lighter model, faster inference speed, but lesser accuracy



Results - Weights and Biases

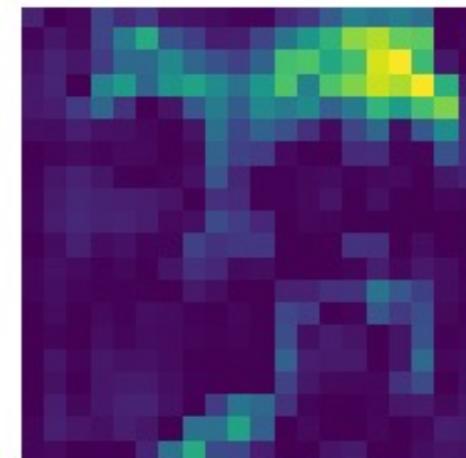
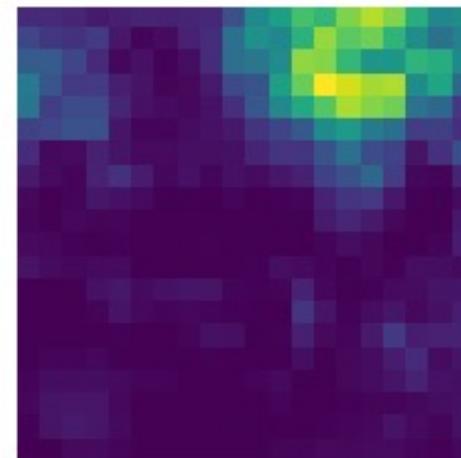
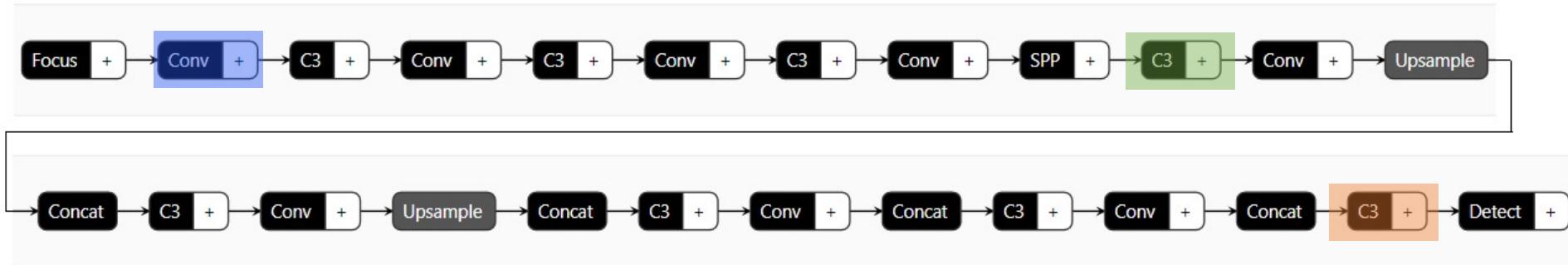


Results



Results - Summary

Visualizing feature maps





Discussion on inference result

GOOD, BAD, AND UGLY...

Ground Truth: maksssksksss253



Inference: maksssksksss253



GOOD

Ground Truth: makssssksksss179



Inference: makssssksksss179



BAD

Ground Truth: maksssskssss633



Inference: maksssskssss633



UGLY



Thanks !