Selected Topics in Visual Recognition using Deep Learning
Homework 4 announcement

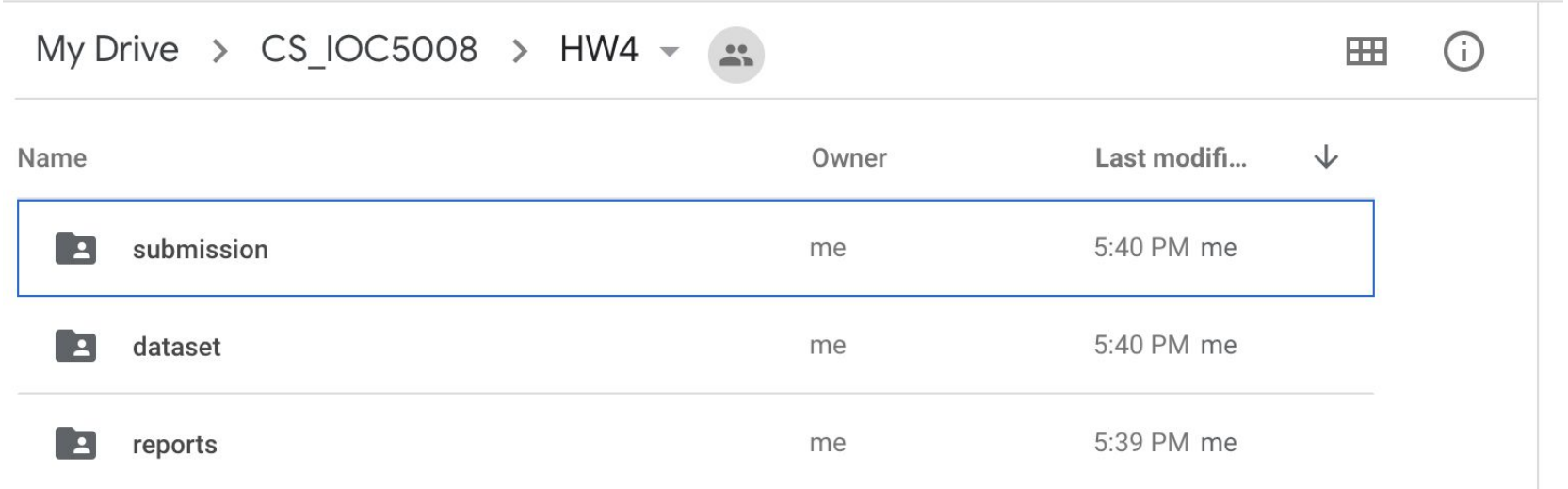
TA: 楊証琨, Jimmy

Ph.D. student at National Taiwan Universitiy

d08922002@ntu.edu.tw

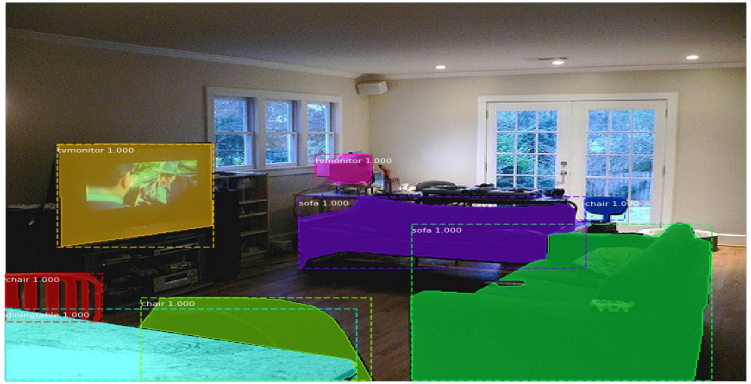
Homework 4: Instance segmentation

- **Deadline: 12/19, Thr at 23:59**
  1. Upload your **report.pdf** and **submission file** in this Google drive



HW4 Introduction: Tiny PASCAL VOC dataset

- Tiny VOC dataset contains only 1,349 training images, 100 test images with 20 common object classes
- **NO external data should be used and only ImageNet pre-trained model can be used**
- Deal with the overfitting problem!



Learn to process PASCAL VOC dataset

- PASCAL VOC datdaset are often evaluated on currenet computer vision models

	mean	aero plane	bicycle	bird	boat	bottle	bus	car
RecoNet152_coco [?]	**89.0**	97.3	**80.4**	96.5	**83.8**	89.5	97.6	95.4
DeepLabv3+_JFT [?]	**89.0**	**97.5**	77.9	96.2	80.4	**90.8**	**98.3**	**95.5**
SRC-B-MachineLearningLab [?]	88.5	97.2	78.6	**97.1**	80.6	89.7	97.4	93.7
DeepLabv3+_AASPP [?]	88.5	97.4	80.3	**97.1**	80.1	89.3	97.4	94.1
SepaNet [?]	88.3	97.2	80.2	96.2	80.0	89.2	97.3	94.7
EMANet152 [?]	88.2	96.8	79.4	96.0	83.6	88.1	97.1	95.0
MSCI [?]	88.0	96.8	76.8	97.0	80.6	89.3	97.4	93.8
ExFuse [?]	87.9	96.8	80.3	97.0	82.5	87.8	96.3	92.6
DeepLabv3+ [?]	87.8	97.0	77.1	**97.1**	79.3	89.3	97.4	93.2
CFNet [?]	87.2	96.7	79.7	94.3	78.4	83.0	97.7	91.6
DeepLabv3-JFT [?]	86.9	96.9	73.2	95.5	78.4	86.5	96.8	90.3

HW4 Get the dataset

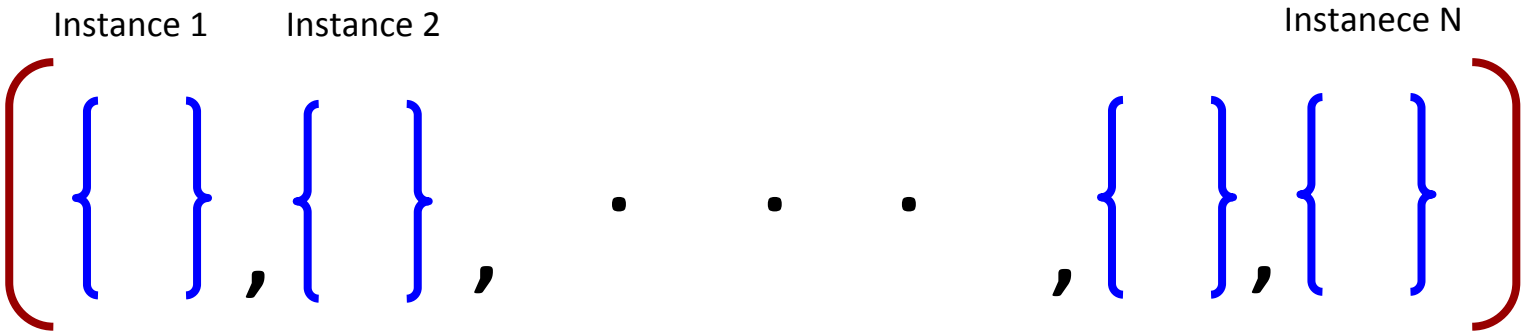
- Download the dataset from this Google Drive
- The annotations are saved in **json file**. You can use pycocotools to read this file
- See data_loader.ipynb for more details

Upload your submission.json file here

- Free version Kaggle doesn't provide the metrics :(
- Upload you submission file into the Google Drive. I will evaluate and return the performance on your filename every midnight
- filename should be STUDENTID.json

Submission.json file format

- List of dictionaries, len(list) = number of detected instance in all test images. Find pseudo code here in *Prepare submission file*
- Each dictionary contains three keys
  - **“image_id”**: id of test image, which is the key in “test.json”, ***int***
  - **“score”**: probability for the class of this instance, ***float***
  - **“category_id”**: category id of this instance, ***int***
  - **“segmentation”**: Encode the mask in RLE by provide function, ***str***



Evaluation metrics: mean Average Precision

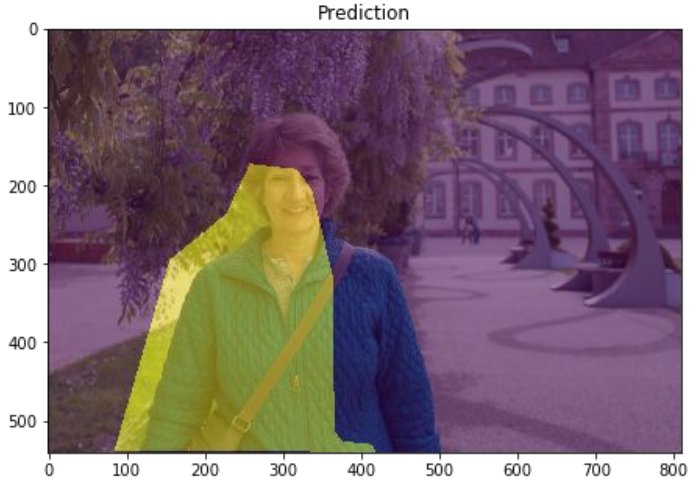
- Most common metric for object detection/segmentation
- Measure the average precision on different threshold and also the IOU between GT and prediction

	backbone	cascade	AP	AP_{50}	AP_{75}
Faster R-CNN	AlexNet	✗	29.4	63.2	23.7
		✓	38.9	66.5	40.5
Faster R-CNN	VGG	✗	42.9	76.4	44.1
		✓	51.2	79.1	56.3
R-FCN	RetNet-50	✗	44.8	77.5	46.8
		✓	51.8	78.5	57.1
R-FCN	ResNet-101	✗	49.4	79.8	53.2
		✓	54.2	79.6	59.2

TABLE 11: Detection results on PASCAL VOC 2007 `test`.

Evaluation metrics: mean Average Precision

- Use mask to measure the Intersection-of-Union of predictions and ground truth
- We use average precision at IOU=0.5 to evaluate your results



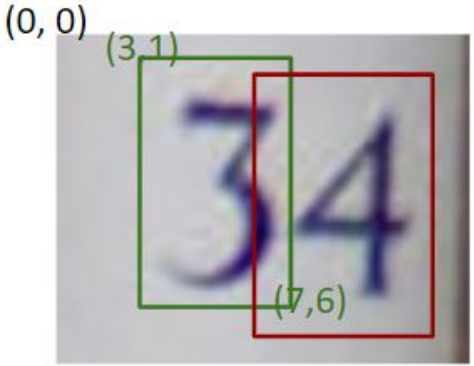
Evaluation metrics: mean Average Precision

- Set IOU threshold=0.5
- We got 1 TP, 1 FP, precision=0.5, recall=0.5
- But if we change the probability threshold to 0.7, the red box will disappear, then we got 1 TP, precision=1, recall=0.5
- We also miss one GT, so when recall=1, precision=0.5

- E.g., your model output two boxes on image, the dictionary will be

y1,x1, y2, x2

```
dict = {"bbox": [(1, 3, 6, 7), (4, 5, 8, 12)] ,
        "label": [3, 5],
        "score":[0.87, 0.61]}
```

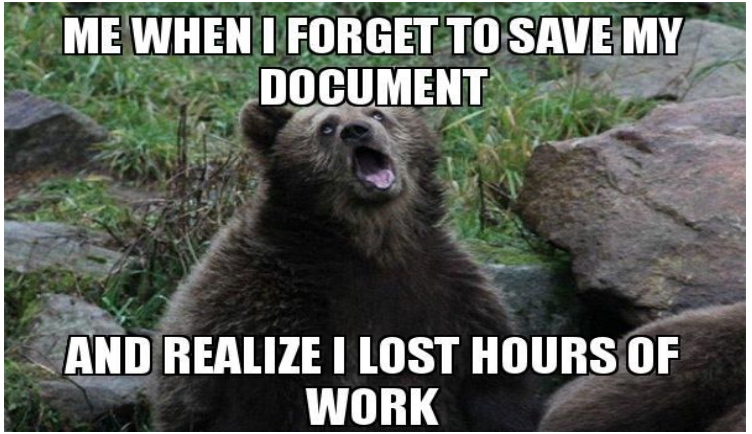


Grading policy: Model performance (70 points)

- Get at least 56% (70%x0.8) by scoring over the baseline
- baseline (mAP@0.5): 0.247

Grading policy: Reports (20 points)

- Document your work (**in PDF**)
  - GitHub/ GitLab link of your code
  - **reference if you used code from GitHub**
  - Brief introduction
  - Methodology (Data pre-process, Model architecture, Hyperparameters,...)
  - Findings or Summary



Grading policy: Code readability (10 points)

- Write beautiful Python code with PEP8 guidelines for readability. Base requirement: use whitespace correctly!

Python

```python
# Recommended
def function(default_parameter=5):
    # ...


# Not recommended
def function(default_parameter = 5):
    # ...
```

Python

```python
# Recommended
my_list = [1, 2, 3]

# Not recommended
my_list = [ 1, 2, 3, ]
```

Python

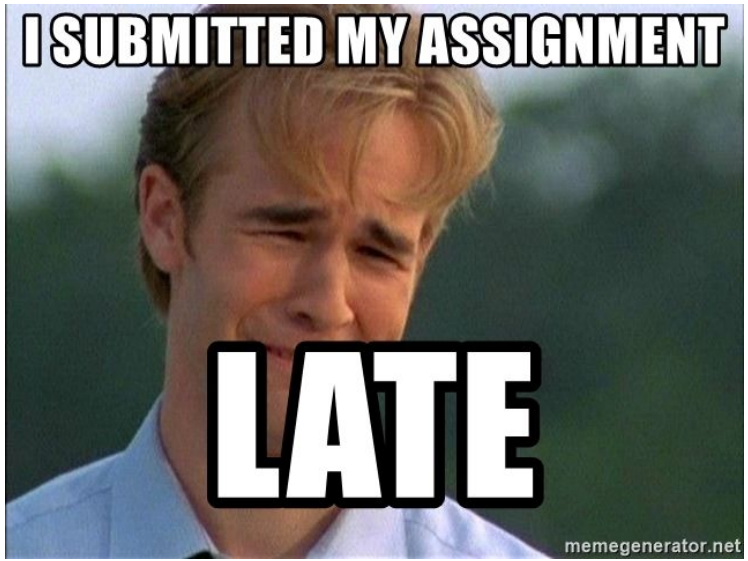
```python
x = 5
y = 6

# Recommended
print(x, y)

# Not recommended
print(x , y)
```

Late Policy

- We will deduct a late penalty of 20 points per additional late day
- For example, If you get 90% of HW but delay for two days, your will get only 90 points- (20 points x 2) = 50 points!



Keywords

- Beat the baseline
  - Mask R-CNN
  - Data-augmentation
  - Regularization (Dropout, L2-norm,...)

- Rank Top 3!
  - Read some paper from CVPR2019, ICCV2019 and try to implement it!

FAQ

- Can I use any code/tools/Library from GitHub or other resources?
  - Yes! We encourage you to learn how to apply existing tools on your own task, such as Keras-Mask R-CNN, Pytorch-maskrcnn-benchmark, TF-object-detection-API

    But DO NOT copy code from your classmate!
- Why my testing results are so bad?
  - CNN model prone to overfitting with small dataset. Use some techniques such as regularization, data-augmentation to solve it!

Notice

- Check your email regularly, we will mail you if there are any updates or problems of the homework
- If you have any questions or comments for the homework, please mail me and cc Prof. Lin
  - Prof. Lin: lin@cs.nctu.edu.tw
  - Jimmy: d08922002@ntu.edu.tw

Have fun!