

Homework #2

Long-Tailed Object Detection



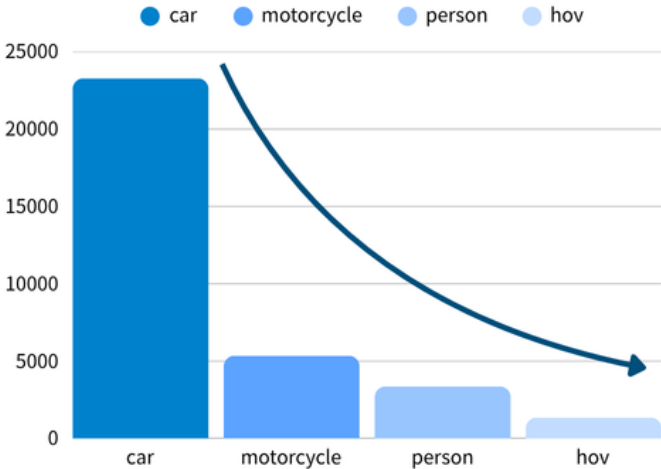
Wen-Huang Cheng (鄭文皇)

National Taiwan University

wenhuang@csie.ntu.edu.tw

TOPIC: Long-tailed object detection for drone-based intelligent counting

- Input: 2D RGB image
- Class label: [car, hov, person, motorcycle]
- Task: long-tailed object detection



2025/10/31 23:59

Kaggle: Hard deadline, No extensions

NTU COOL: Late Submission Policy is Applicable

(See TA01-Assignments_Overview.pdf)

- **We have separately created Kaggle competitions exclusively for NTU students and exclusively for TAICA**
- **Please only participate in the competition designated for you (NTU or TAICA)**
- **The grading policy for TAICA students should be determined by your institution's teaching assistants or professors. Any questions, please contact them.**



Dataset Description

- Download link: On Kaggle competition
- Training set: 950 images with 33,331 objects. Each image has a corresponding gt.txt file (e.g. img0001.png, img0001.txt).

Each line in gt.txt contains:

<class label>,<Top-left X>,<Top-left Y>,<Bounding box width>,<Bounding box height>

- Testing set: 550 images with 27315 objects. (Image only.)
- **DO NOT** try to find the ground truth of testing set
- **DO NOT** use testing data during training
 - Violating the rules on this page will result in a score of zero.
 - If you are uncertain about the legitimacy of the usage, email the TAs for clarification

- Average Precision $mAP_{50:95}$
 - The average of the mean average precision calculated at varying IoU thresholds, ranging from 0.50 to 0.95.
 - Please refer to the course slides or this [intro](#).
- There are one simple baseline and one strong baseline, beat them to get the higher score.



Grading Policies

- Kaggle Competition (testing set) (90%) (according to **private leaderboard**)

μ, σ are calculated on $\text{mAP}_{50:95} \geq \text{private baseline}$.

Your $\text{mAP}_{50:95}$	Points
$\text{mAP}_{50:95} \geq \mu + 2\sigma$	90
$\text{mAP}_{50:95} \geq \mu + \sigma$	80
$\text{mAP}_{50:95} \geq \mu$	75
$\text{mAP}_{50:95} \geq \mu - \sigma$	70
$\text{mAP}_{50:95} \geq \mu - 2\sigma$	65
$\text{mAP}_{50:95} \geq \text{strong baseline}$	60
$\text{mAP}_{50:95} \geq \text{simple baseline}$	Linear between 50 ~ 60

- Report (10%)



Kaggle (90%) (For **all** Students)

- This is a private competition. Please join using the email you use for NTU COOL. If you find that you cannot join the competition, please email the TAs to request being added.
- The maximum daily submissions is 5.
- The public leaderboard shows the score of only 50% test data. Your final score is evaluated on the other 50% test data.
- You can optionally select 2 submissions as the final submissions.
- **DO NOT** use more than one account to participate in the Kaggle competition.



Kaggle (90%) (For **NTU** Students)

- Use this [link](#) to participate in the competition
- **DO** rename your team name to <student-id> (e.g., R12345678).
- **DO NOT** use more than one account to participate in the Kaggle competition.
 - NTU students should not participate in the TAIKA competition.
 - Violating the rules on this page will result in a score of zero.



Kaggle (90%) (For **TAICA** Students)

- Use this [link](#) to participate in the competition.
- **DO** rename your team name to <**school**_student-id> (e.g., **NYCU_R12345678**).
- **DO NOT** use more than one account to participate in the Kaggle competition.
 - **TAICA students should not participate in the NTU competition.**
 - **Violating the rules on this page will result in a score of zero.**



Kaggle Submission Format

- Save predictions in csv format and submit to Kaggle:

```
Image_ID,PredictionString
1,<conf_1> <bb_left_1> <bb_top_1> <bb_width_1> <bb_height_1> <class_1> <conf_2>...
2,<conf_1> <bb_left_1> <bb_top_1> <bb_width_1> <bb_height_1> <class_1> <conf_2>...
...
```

- Image_ID is the index of images in the test set.
- PredictionString contains ALL DENORMALIZED predictions of the corresponding image.
- Use “,” to separate Image_ID and PredictionString **WITHOUT SPACE**.
- Use “**SPACE**” to separate the attributes of bounding boxes.
- The mapping between class names and class labels is as follows:
0: car, 1: hov, 2: person, 3: motorcycle



Programming Spec

- Use Python ≥ 3.8 (for consistency and reproducibility).
- No pretrained weights can be used for any component, including encoders.
Model choice is unrestricted, but total VRAM usage must not exceed 12 GB.
- Only use the provided dataset — no extra datasets allowed.
- No plagiarism (Online resources may be consulted. However, using the same code source as classmates may be considered plagiarism).
- Violating the above rules on this page will result in a score of zero.
- If you are uncertain about the legitimacy of the usage, email the TAs for clarification

1. Model Description

- Introduce your model (must include an **architecture illustration & any modifications**)

2. Implementation Details

- Preprocessing, augmentation, hyperparameters, loss functions, training strategies, etc.

3. Result Analysis

- Quantitative improvements (tables, metrics, discussion)
- Visualizations (e.g., example detections, error analysis)
- Analyze how your method addresses the long-tailed problem, providing comparative results before and after. You may present evidence of effectiveness in multiple ways.

4. Short conclusion

- 3-5 pages (excluding references), exceeding the limit will result in a -5 score

- Your submission should be a zipped file with the following structure:
 - hw2_<student-id>.zip (e.g., hw2_R12345678.zip)
 - |-- hw2_<student-id> (Should contain this folder, not separate files)
 - |----- report_<student-id>.pdf (Your report)
 - |----- code_< student-id>.zip
 - |----- src/ (Your source code)
 - |----- readme.md (Show how to install the environment, run training & prediction)
 - |----- requirements.txt (The list of necessary packages)
- An incorrect format will result in a deduction of a -5 score.
- Failure to re-implement similar performance will result in a 60% discount of the total score.
- Plagiarism in the report or code will result in 0%.

Any Question

Ask peers first

(Join with name: <school_student-id>)



Then ask TAs

(only for NTU students)

cvpdl.ta.fall.2025@gmail.com