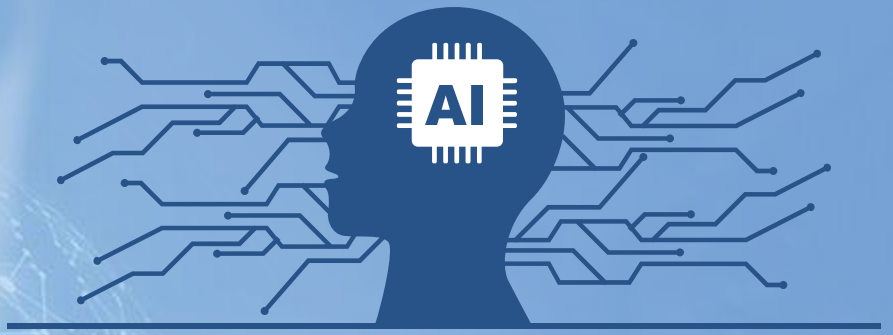
CVPDL: Computer Vision Practice with Deep Learning



Homework #1
Object Detection

Wen-Huang Cheng (鄭文皇)

National Taiwan University

wenhuang@csie.ntu.edu.tw

HW1 - Object Detection

TOPIC: Object Detection for Group-hosed Swine

- Input: 2D RGB image
- Task: object detection

Submission Deadline

2025/10/10 23:59

Kaggle: Hard deadline, No extensions

NTU COOL: Late Submission Policy is Applicable

(See TA01-Assignments_Overview.pdf)

NTU & TAICA

- **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: 1266 images (including img/ & gt.txt)

  Each line in gt.txt contains:

```
<frame>, <bb_left>, <bb_top>, <bb_width>, <bb_height>
```

- Testing set: 1864 images (including img/)
- 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

Evaluation Metric

- Average Precision $\mathrm{mAP}_{50:95}$
  - Since we only have one class, it equals to $\mathrm{AP}_{50:95}$.
  - 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 $\text{mAP}_{50:95} \geq$ 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
$\text{mAP}_{50:95} \geq$ strong baseline	60
$\text{mAP}_{50:95} \geq$ simple baseline	Linear between 50 ~ 60

- Report (10%)

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

- ➢ Use this link to participate in the competition.
- ➢ DO rename your team name to <student-id> (e.g., R12345678).
- ➢ 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.

- Violating the rules on this page will result in a score of zero.
- Students of TAICA DO NOT participate in this competition

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).
- 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.

- Violating the rules on this page will result in a score of zero.
- Students of NTU DO NOT participate in this competition

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.
- Set <class> to **0** to represent pigs.

Programming Spec

- ➢ Use Python >= 3.10 (for consistency and reproducibility).
- ➢ No loading of pretrained weights (except as feature extractors).
- ➢ 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

Report (10%)

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)
4. **Short conclusion**

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

NTU COOL Submission Rules

➢ Your submission should be a zipped file with the following structure:

- hw1_<student-id>.zip (e.g., hw1_R12345678.zip)

  |-- hw1_<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

Supplementary Material

Download Dataset to Google Drive on Colab

- Kaggle → your profile → Account → Create New API token → upload token to google drive ./.kaggle/kaggle.json.
- Mount google drive to /content/drive.

```
from google.colab import drive
drive.mount('/content/drive')
```

- Use kaggle package to download dataset.

For NTU students

```
! KAGGLE_CONFIGDIR=/content/drive/MyDrive/.kaggle \
kaggle competitions download \
-c ntu-cvpdl-2025-hw-1\
-p ./drive/MyDrive
```

For TAICA students

```
! KAGGLE_CONFIGDIR=/content/drive/MyDrive/.kaggle \
kaggle competitions download \
-c taica-cvpdl-2025-hw-1\
-p ./drive/MyDrive
```

Download Dataset to Google Drive on Colab

➢Copy dataset to VM storage and unzip it.

For NTU students

```
! cp ./drive/MyDrive/ntu-cvpdl-2025-hw-1.zip .
! unzip ./ntu-cvpdl-2025-hw-1.zip -d ./data
```

For TAICA students

```
! cp ./drive/MyDrive/taica-cvpdl-2025-hw-1.zip .
! unzip ./taica-cvpdl-2025-hw-1.zip -d ./data
```

➢Start to read data from path ./data/train and ./data/test

Requirements.txt

➢Generate requirements.txt

```
pip freeze > requirements.txt
```

➢Inatall package from requirements.txt

```
pip install -r requirements.txt
```

➢Manually create requirements.txt

```
pandas
numpy==2.2.6
torch==2.4.0
torchvision==0.19.0
...
```

Use Virtual Environment

➢Create virtual environment for a project and install the needed packages only, for example:

- virtualenv
- conda
- pyenv