Selected Topics in Visual Recognition using Deep Learning
Homework 4 announcement

TA: 楊証琨, Jimmy

Ph.D. student at National Taiwan Universitiy

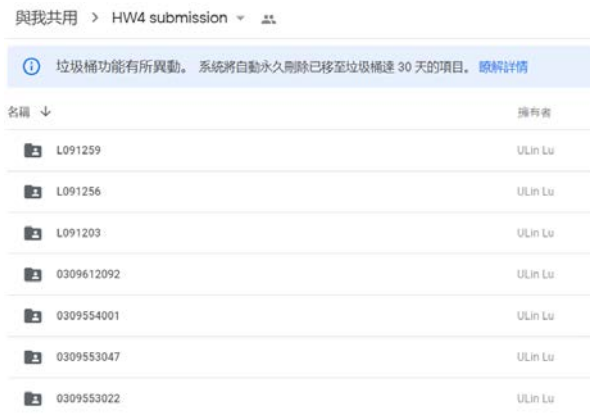
d08922002@ntu.edu.tw

Homework 3 reminder

- **Deadline: 12/10** (one week left)
  - Upload your **submission file** in this Google drive
  - Upload your **report.pdf** to E3
- **DO NOT** use the COCO/VOC pre-trained model and make sure your result is reproducible. Otherwise, you will get only half points of the model performance

Homework 4: Image super-resolution

- **Deadline: 01/07, Thr at 23:59** (You have 5 weeks to compelete HW4)
  1. Upload your **report.pdf** to E3
  2. Upload your **upscaled images** into this Google drive
- **DO NOT** use any pre-trained models or external data



HW4 Introduction: Image super resolution

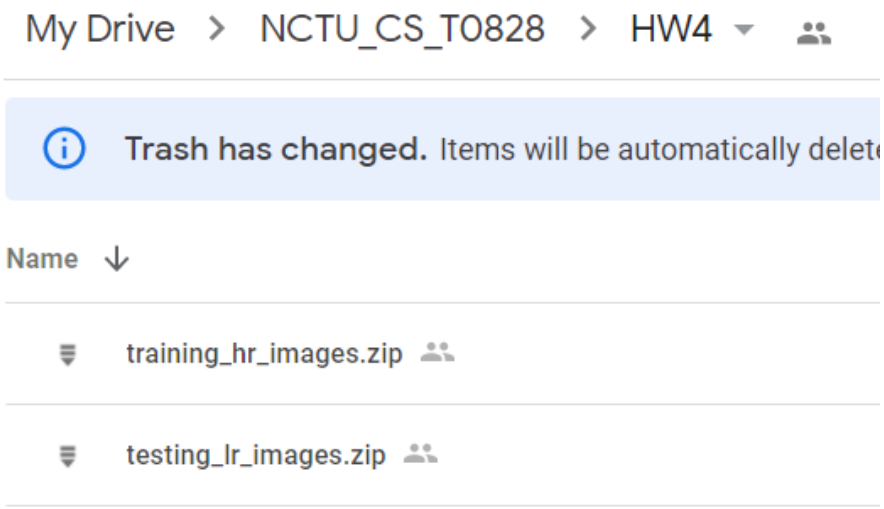
- Dataset
  - Training set: 291 high-resolution images
  - Testing set: 14 low-resolution images
- Train your model to reconstruct a high-resolution image from a low-resolution input

low-resolution image

high-resolution image

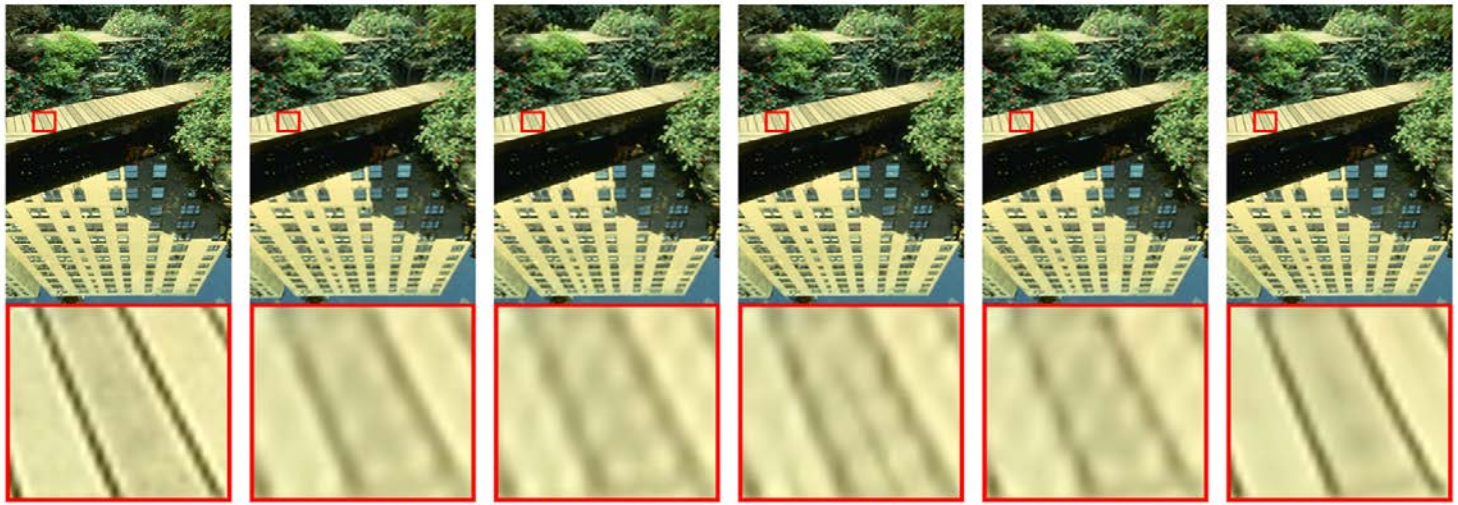
HW4 Get the dataset

- Download the dataset from this Google Drive
- There is no annotations in the image super-resolution task. Create the HR-LR image pairs by the provided HR images



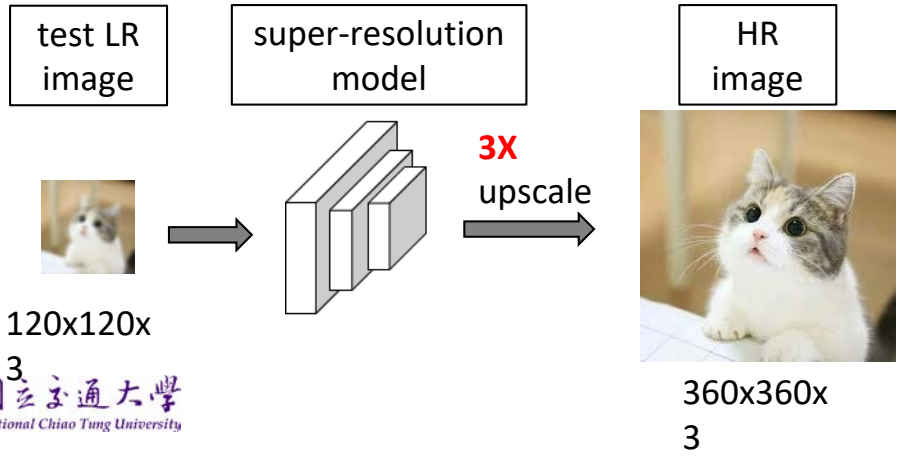
Evaluation metrics: PSNR

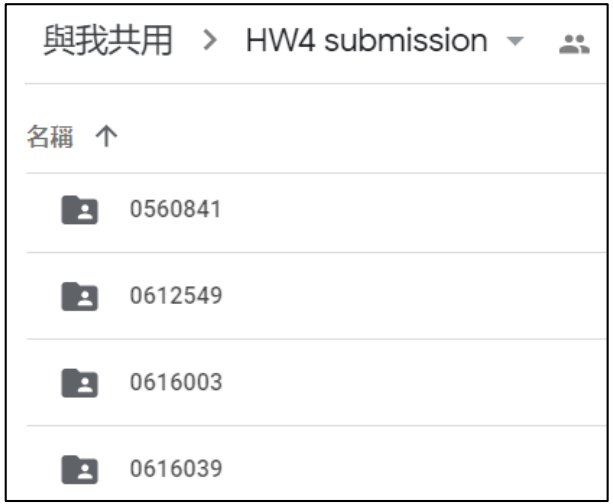
- Peak signal to noise ratio (PSNR) is a commonly used metric to measure the similarity between two image
- Baseline performance in PSNR: **29.77**



Upload your upscaled images here

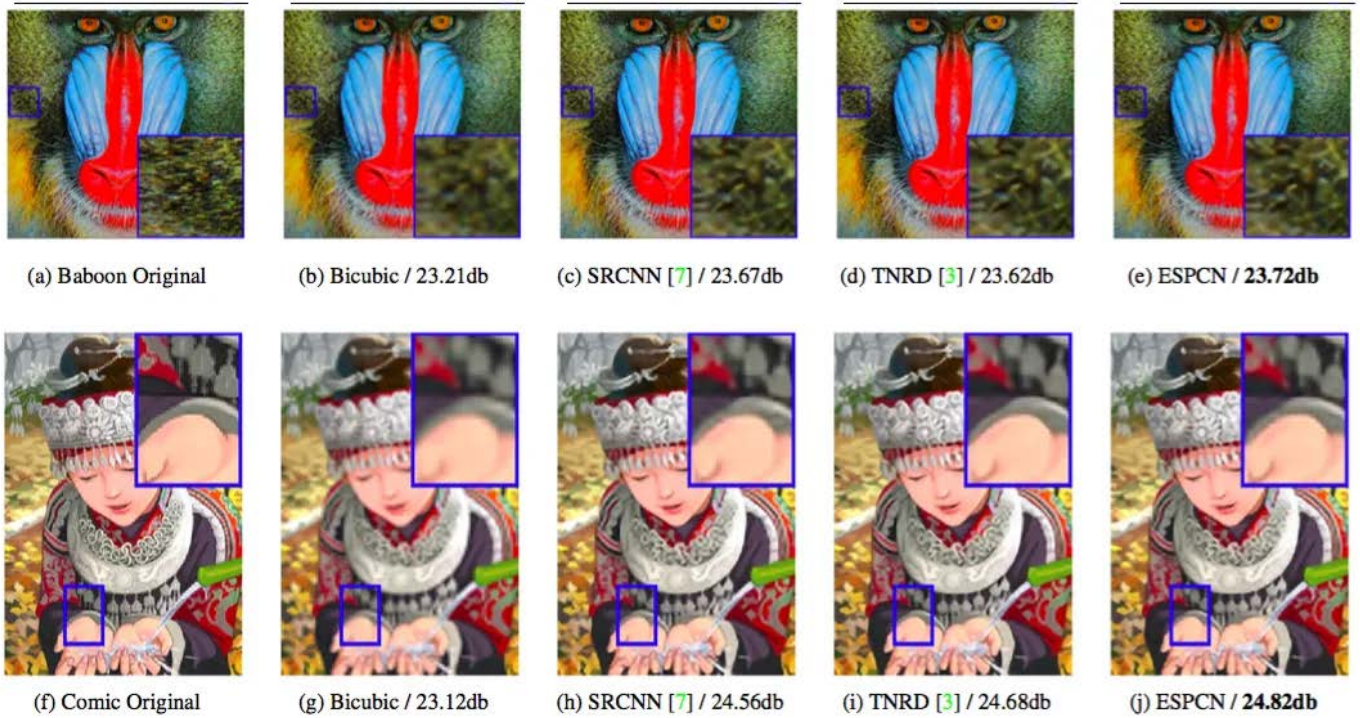
- Please upscale the test images with an upscaling factor of **3**, e.g., 120x120 -> 360x360
- Upload your generated high-resolution images into your student ID folder. We will return the PSNR results periodically on your folder name





Grading policy: Model performance (70 points)

- Get at least 56% (70%x0.8) by scoring over the baseline



(a) Baboon Original (b) Bicubic / 23.21db (c) SRCNN [7] / 23.67db (d) TNRD [3] / 23.62db (e) ESPCN / **23.72db**

(f) Comic Original (g) Bicubic / 23.12db (h) SRCNN [7] / 24.56db (i) TNRD [3] / 24.68db (j) ESPCN / **24.82db**

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

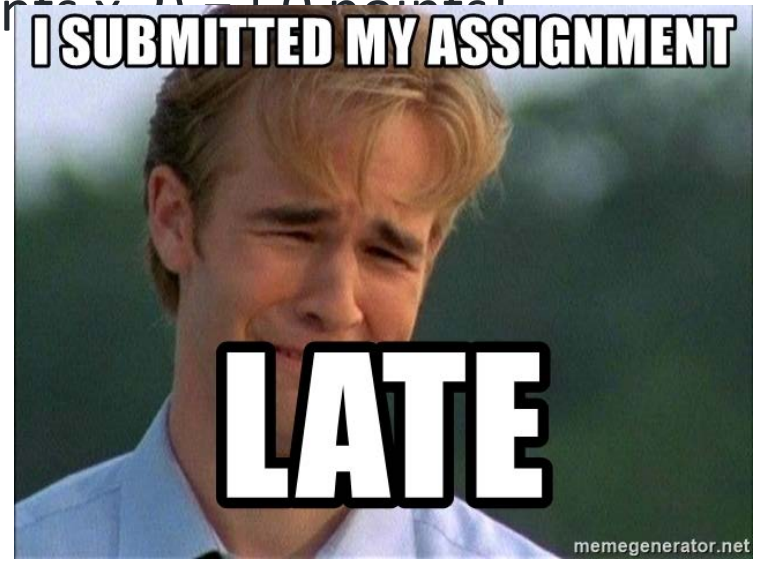
```python
x = 5
y = 6

# 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
  - VDSR [Kim etal. CVPR'16]
  - Data-augmentation

- Rank Top 3!
  - Read the SOTA paper (image super-resolution) from PAPERs-with-codes 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.

    But DO NOT copy code from your classmate!
- Why my testing results are so bad?
  - **DO NOT** rename the test images when uploading your results
  - Please upscale the test images with an scaling factor of **3**!

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** (3pm-4pm, Thur., EC118)
  - 佳諭: **mylifeai1116@gmail.com** (3pm-4pm, Thur., EC118)
  - 玉霖: **oscar861201@gmail.com** (3pm-4pm, Thur., EC118)

Have fun!

Selected Topics in Visual Recognition using Deep Learning
Final project announcement

TA: 楊証琨, Jimmy

Ph.D. student at National Taiwan Universitiy

d08922002@ntu.edu.tw

Final project: Join a competition on Kaggle

- Team up
  - **3 persons per team.** Please fill this google forms by your team leader before 12/5, Sat at 23:59
  - We will randomly match up if you can not form a team after 12/5, Sat (If you are planned to withdraw this class, please inform Prof. Lin and all TAs by 12/5)
- Select a competition from 4 candidates
  - Complete this google forms by your team leader before 12/10, Thr at 23:59
- Presentation and report submission: 12/24, 12/31, Thr

Final project: Join a competition on Kaggle

1. Join one of the provided competitions and beat the baseline
2. Make a 10-min presentation in English of your methodology and have a 2-min Q&A session
3. Upload your report and slides for a team (one report for each team)

Baseline

Competition Medals

Competition medals are awarded for top competition results. The number of medals awarded per competition varies depending on the size of the competition. Note that InClass, playground, and getting started competitions do not award medals.

	0-99 Teams	100-249 Teams	250-999 Teams	1000+ Teams
Bronze	Top 40%	Top 40%	Top 100	Top 10%
Silver	Top 20%	Top 20%	Top 50	Top 5%
Gold	Top 10%	Top 10	Top 10 + 0.2%*	Top 10 + 0.2%*

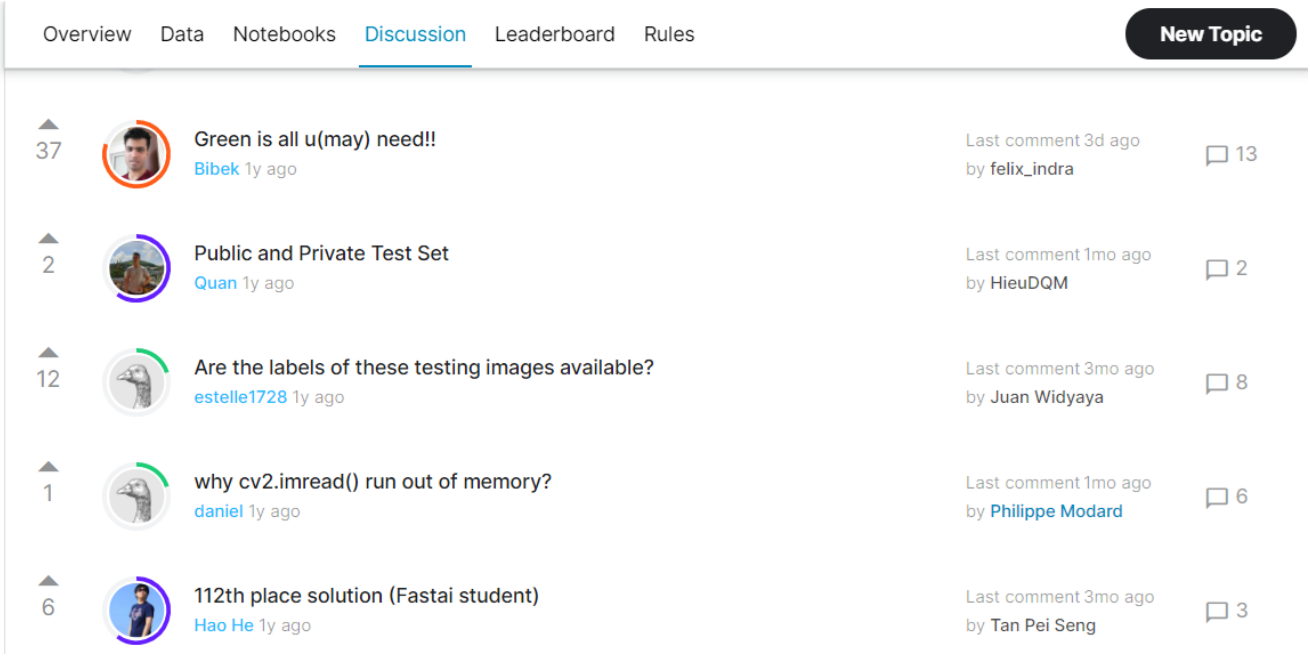
Competitions on Kaggle

1. APTOS 2019 Blindness Detection: Image classficiation
2. Handwritten Grapheme Classification: Image classficiation
3. Global Wheat Detection: Object detection
4. Severstal: Steel Defect Detection: Semantic segmentation

- **Report & slides submission before the presentation**
- Present in this order (sorted by difficulty). You will present earlier if selecting the easier competition. For example, if you select APTOS, your presentation is probably on 12/24.

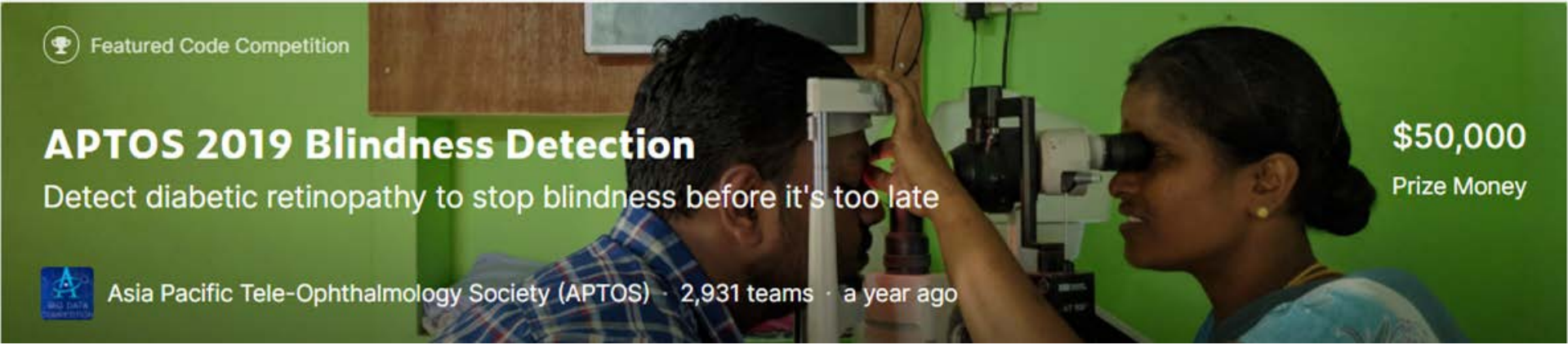
Learn from the competitors

- Find some useful insights in the dicussion



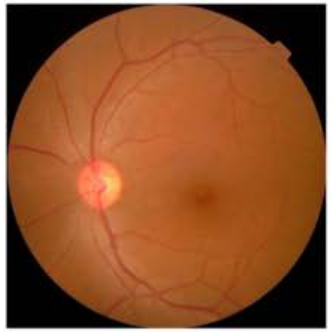
1. APTOS 2019 Blindness Detection

- Dataset: 3,662 images, 8 GB, 5 class classification
- Evaluation: quadratic weighted kappa, which measures the agreement between two ratings
- Baseline: silver medal (top 5% of 2,931 teams)

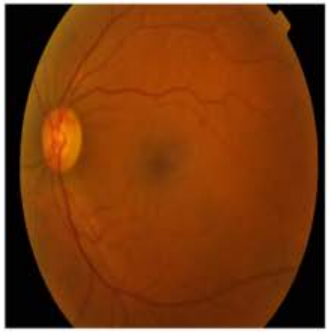


1. APTOS 2019 Blindness Detection

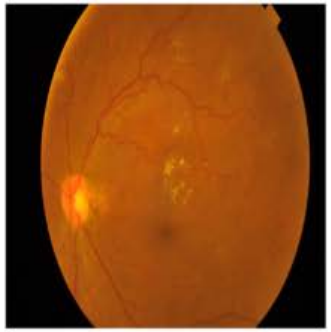
- Classify the **ratings** correctly for all images



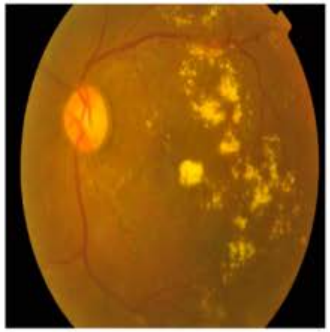
Normal



Mild



Moderate



Severe

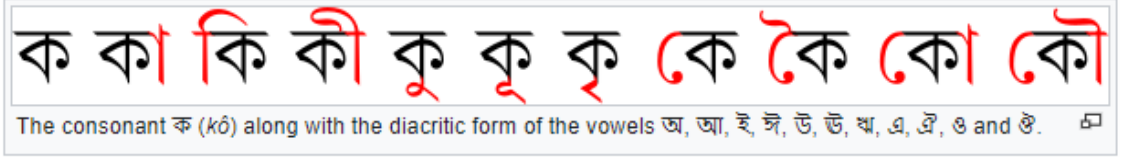
Proliferative

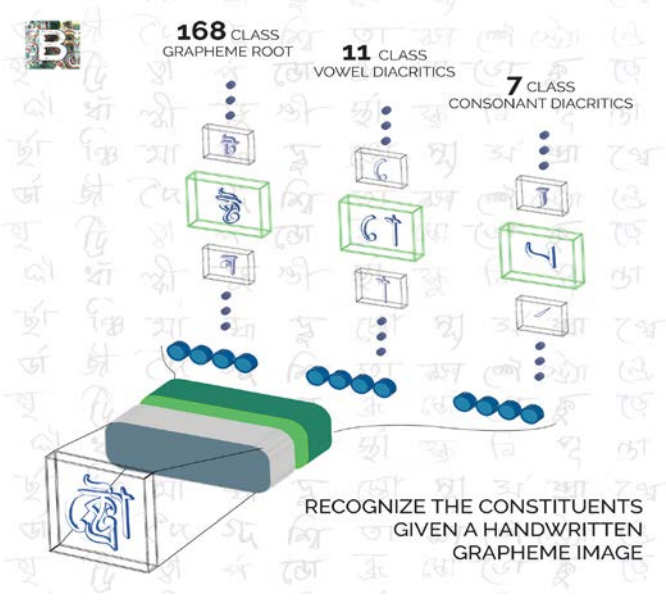
2. Bengali.AI Handwritten Grapheme Classification

- Dataset: 411,882 images (137x263), 4.8 GB
- Evaluation: hierarchical macro-averaged recall
- Baseline: silver medal (top 5% of 2,059 teams)

2. Bengali.AI Handwritten Grapheme Classification

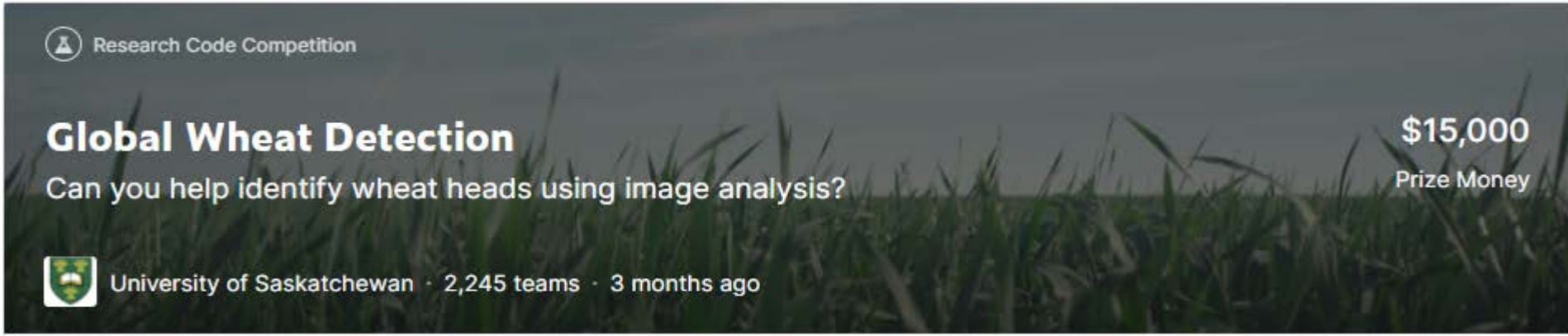
- Classify the **grapheme root** (字根), **vowel diacritic** (母音), and **consonant diacritic** (輔音) for all images
- Your model need to separately classify three constituent elements in the image





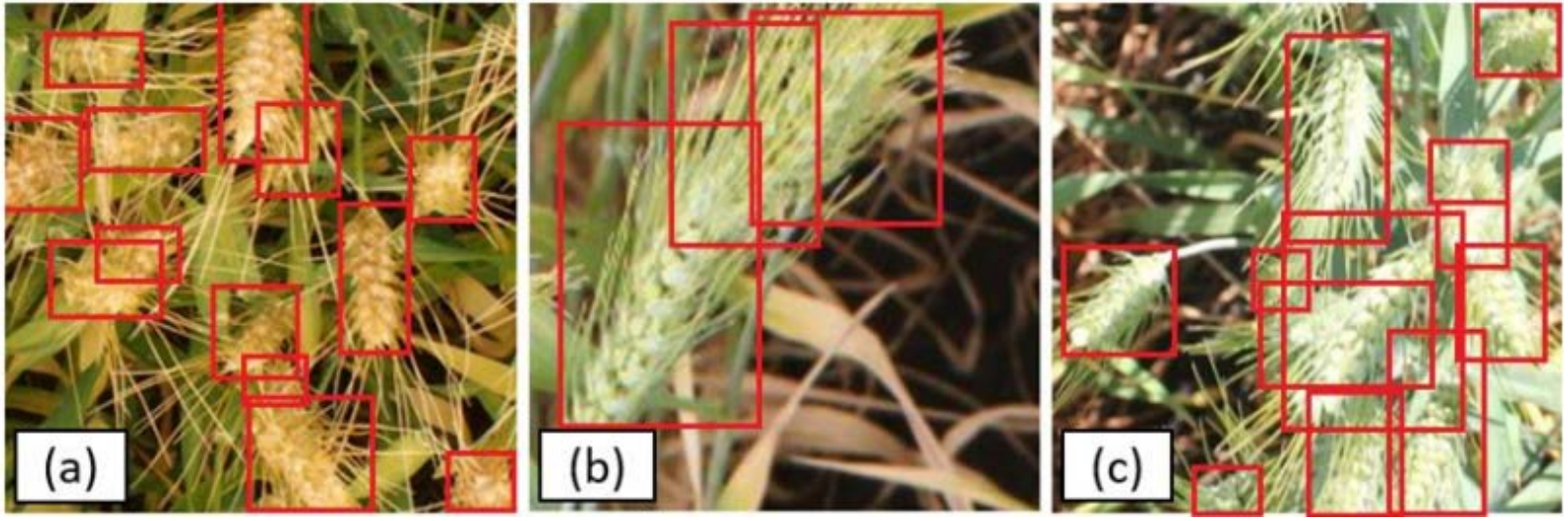
3. Global Wheat Detection

- Dataset: 3,432 images, 613 MB, object detection
- Evaluation: mean average precision
- Baseline: bronze medal (top 10% of 2,245 teams)
- ECCV 2020 workshop



3. Global Wheat Detection

- Detect the wheat heads correctly for all images

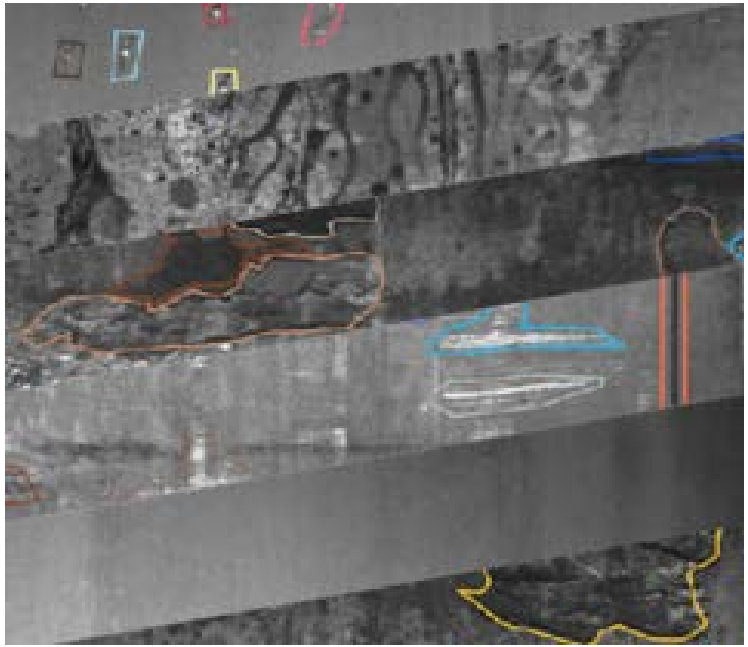


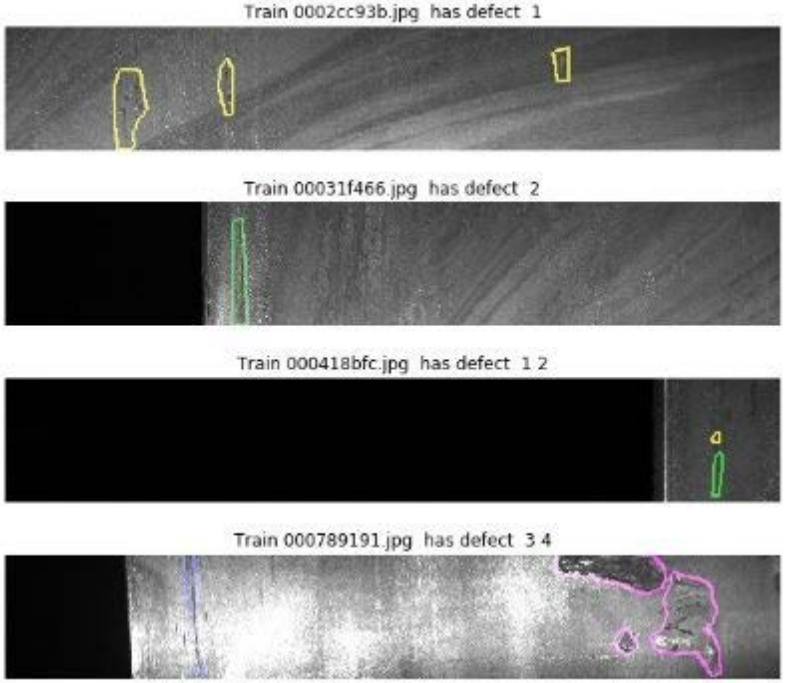
4. Severstal: Steel Defect Detection

- Dataset: 1,800 images, 1.58 GB
- Evaluation: Dice coefficient
- Baseline: bronze medal (top 10% of 2,431 teams)

4. Severstal: Steel Defect Detection

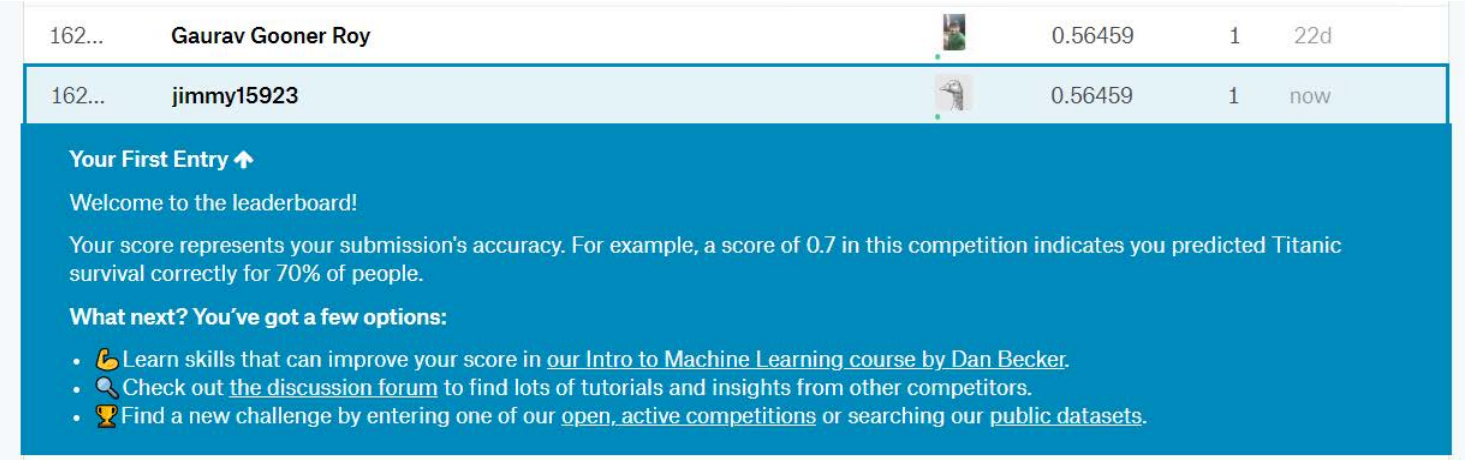
- Segment the **4 types of defect** correctly for all images





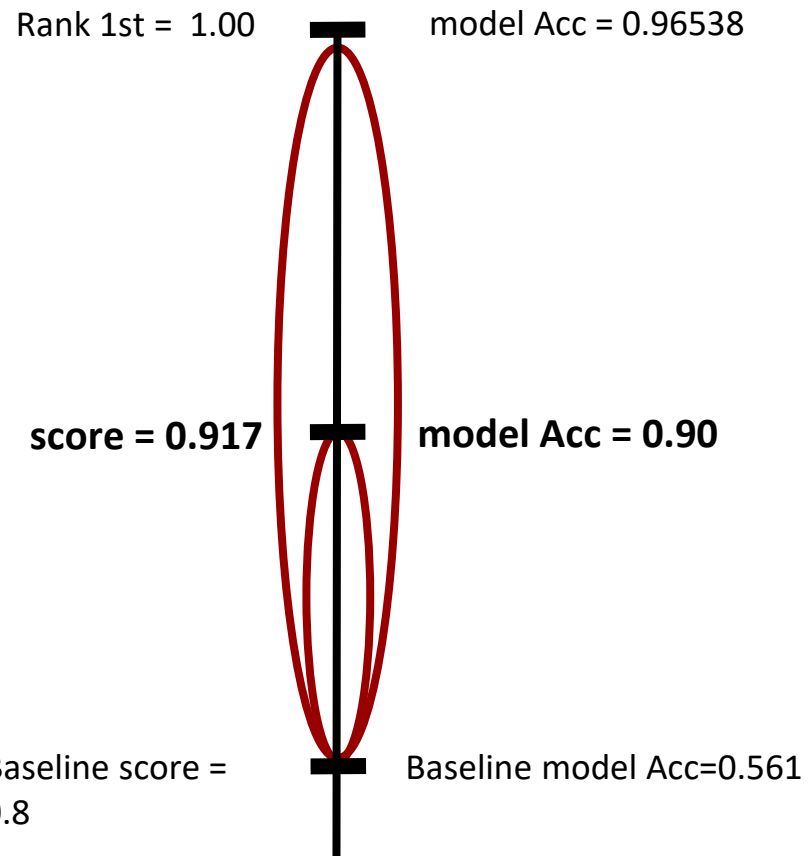
Grading policy

- Model performance: **40 points**
  - ➢ Screenshot your rank and post it in the reports
- Presentation & Report & Code: **50 points**
- Teammate contribution: **10 points**



Grading policy: Model performance (40 points)

- Your score will be interpolated with the model accuracy by rank 1th and baseline



Grading policy: Presentation (50 points)

- Your presentation/reports should include
  - GitHub/ GitLab link of your code
  - Introduction
  - Related work
  - Proposed approach
  - Experiment results
  - Conclusion
- Meet all requirements can get 80% points (40 points)

Grading policy: Presentation - Introduction

- Problem statement
- Importance or value of this problem

Deep Residual Learning for Image Recognition

Kaiming He Xiangyu Zhang Shaoqing Ren Jian Sun

Driven by the significance of depth, a question arises: *Is learning better networks as easy as stacking more layers?* An obstacle to answering this question was the notorious problem of vanishing/exploding gradients [1, 9], which hamper convergence from the beginning. This problem, however, has been largely addressed by normalized initialization [23, 9, 37, 13] and intermediate normalization layers [16], which enable networks with tens of layers to start converging for stochastic gradient descent (SGD) with back-propagation [22].

Grading policy: Presentation - Related work

- Divide the related work into groups
- Summarize the prons and cons of each groups
- Bonus:
  - Identity the advantage of your work over those work in the literature

Focal Loss for Dense Object Detection

Tsung-Yi Lin Priya Goyal Ross Girshick Kaiming He Piotr Dollár

2. Related Work

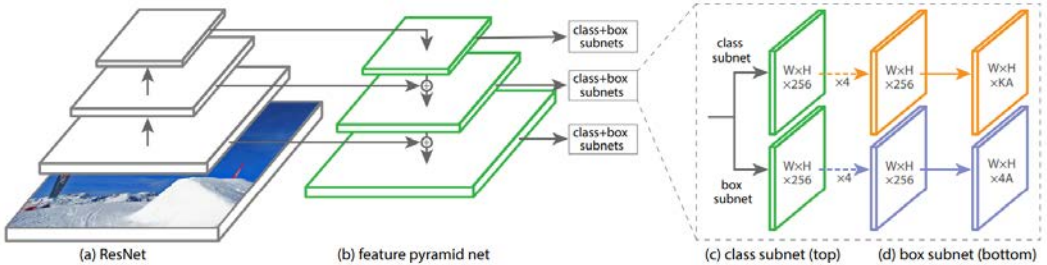
Classic Object Detectors: The sliding-window paradigm, in which a classifier is applied on a dense image grid, has a long and rich history. One of the earliest successes is the

Two-stage Detectors: The dominant object detection is based on a two-stag neered in the Selective Search work [35

One-stage Detectors: OverFeat [30] wa modern one-stage object detector based More recently SSD [22, 9] and YOLO

Grading policy: Presentation - Proposed approach

- Overview: Network figure with loss function
- Details of your approach
- Bonus:
  - ➢ Novelty of your approach over the related work



$$\mathrm{FL}(p_t) = -\alpha_t(1-p_t)^{\gamma}\log(p_t)$$

Grading policy: Presentation - Experimental results

- Evaluation dataset, metrics
- Comparison with SOTA methods (related work) in this competition
- Bonus:
  - Ablation study of the proposed component

	backbone	AP	AP_{50}	AP_{75}	AP_S	AP_M	AP_L
Two-stage methods							
Faster R-CNN+++ [16]	ResNet-101-C4	34.9	55.7	37.4	15.6	38.7	50.9
Faster R-CNN w FPN [20]	ResNet-101-FPN	36.2	59.1	39.0	18.2	39.0	48.2
Faster R-CNN by G-RMI [17]	Inception-ResNet-v2 [34]	34.7	55.5	36.7	13.5	38.1	52.0
Faster R-CNN w TDM [32]	Inception-ResNet-v2-TDM	36.8	57.7	39.2	16.2	39.8	**52.1**
One-stage methods							
YOLOv2 [27]	DarkNet-19 [27]	21.6	44.0	19.2	5.0	22.4	35.5
SSD513 [22, 9]	ResNet-101-SSD	31.2	50.4	33.3	10.2	34.5	49.8
DSSD513 [9]	ResNet-101-DSSD	33.2	53.3	35.2	13.0	35.4	51.1
RetinaNet (ours)	ResNet-101-FPN	39.1	59.1	42.3	21.8	42.7	50.2
RetinaNet (ours)	ResNeXt-101-FPN	**40.8**	**61.1**	**44.1**	**24.1**	**44.2**	51.2

Grading policy: Presentation - Conclusion

- Summarize your work
- Summarize what you learn/found in the final project
- Bonus:
  - Interesting findings/contributions to the task you choose

Slides organization

- Prof. Lin will instruct how to organize yours slides for the 10-min presentation on **12/17, Thr**

Grading policy: Contribution (10 points)

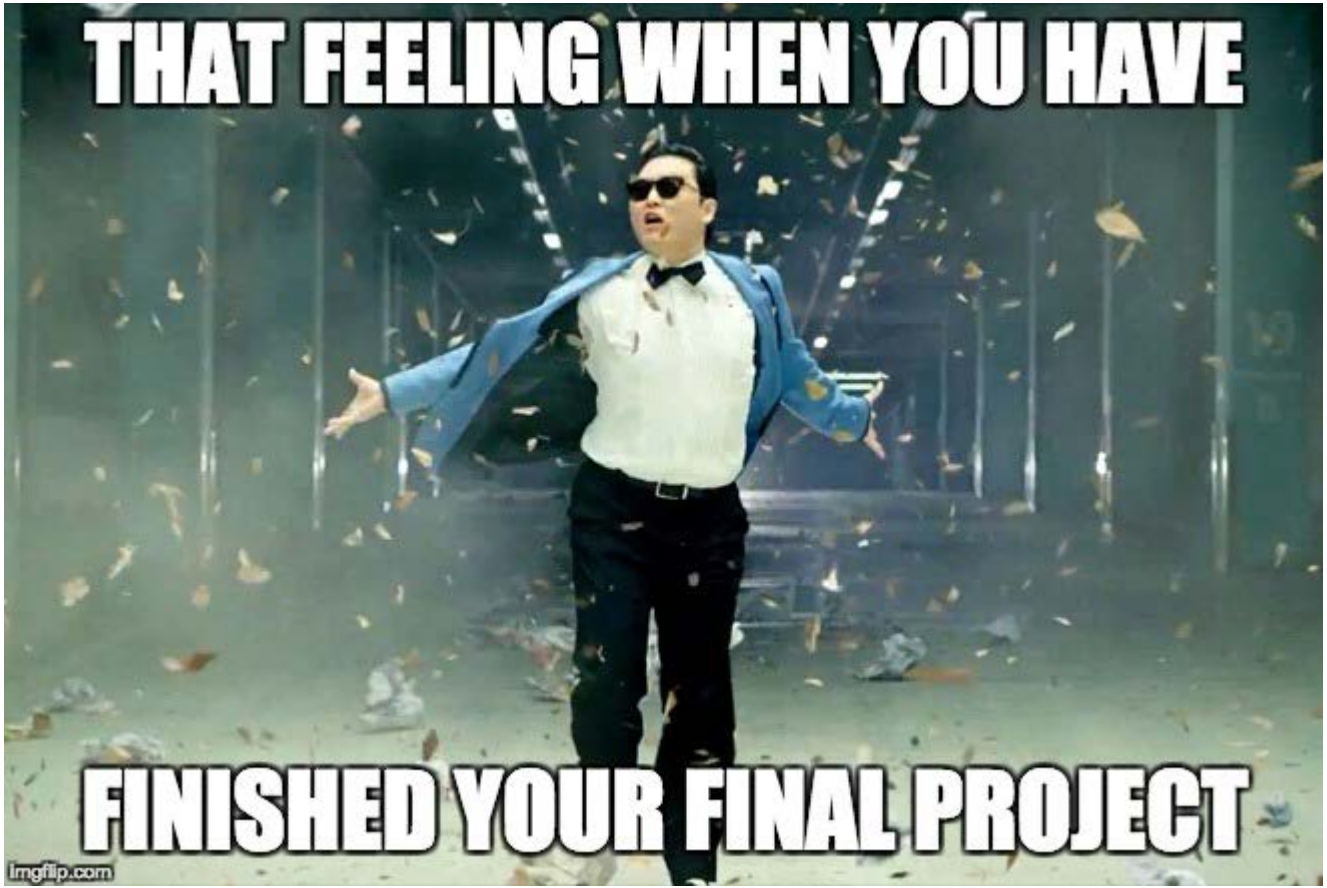
- Specify the teamwork of each tasks from your team **in the reports**

Tasks	**contributors (%)**
Literature survey	0856065 (100%)
Approach design	0856078 (50%), 0856605 (50%)
Approach implementation (experiment)	0856078 (30%), 0856605 (70%)
Report writing	0856065 (80%), 0856078 (20%)
Slide making and oral presentation	0856605 (33%), 0856065 (33%), 0856078 (33%)

Important dates

Event	date
Team up (google form)	**12/5 Sat**
Select competition (google form)	**12/10 Thr**
Final presentation Ⅰ & report submision	**12/24 Thr**
Final presentation Ⅱ & report submision	**12/31 Thr**
Deadline of HW4	**01/07 Thr**

Good luck!


THAT FEELING WHEN YOU HAVE
FINISHED YOUR FINAL PROJECT
imgflip.com