



# Selected Topics in Visual Recognition using Deep Learning

## Homework 4 announcement

TA: 楊証琨, Jimmy  
Ph.D. student at National Taiwan University  
[d08922002@ntu.edu.tw](mailto: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

與我共用 > HW3 submission

垃圾桶功能有所異動，系統將自動永久刪除已移至垃圾桶達 30 天的項目。瞭解詳情

名稱	擁有者	我上次修改的時間	檔案大小
mAP_0.70602_0856566.json	阮建元		745 KB
mAP_0.67975_309551033.json	歐陽美		137 KB
mAP_0.62185_0856095.2.json	黃怡潔		503 KB
mAP_0.58191_0856724.1.json	王彥霖		331 KB
mAP_0.55927_0856724.2.json	王彥霖		430 KB
mAP_0.54006_0856724.json	王彥霖		209 KB
mAP_0.38246_0856610.json	郭聖弘		996 KB
mAP_0.34804_309551033.json	沈欣瑜		129 KB
mAP_0.34416_309551048.json	陳日達		273 KB
mAP_0.24672_BASELINE_submission_sample.json	Ulin Lu		104 KB





# Homework 4: Image super-resolution

- **Deadline: 01/07, Thr at 23:59** (You have 5 weeks to complete 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

與我共用 > HW4 submission ▾

① 垃圾桶功能有所異動。系統將自動永久刪除已移至垃圾桶達 30 天的項目。瞭解詳情

名稱 ▾	擁有者
L091259	ULin Lu
L091256	ULin Lu
L091203	ULin Lu
0309612092	ULin Lu
0309554001	ULin Lu
0309553047	ULin Lu
0309553022	ULin Lu





# 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

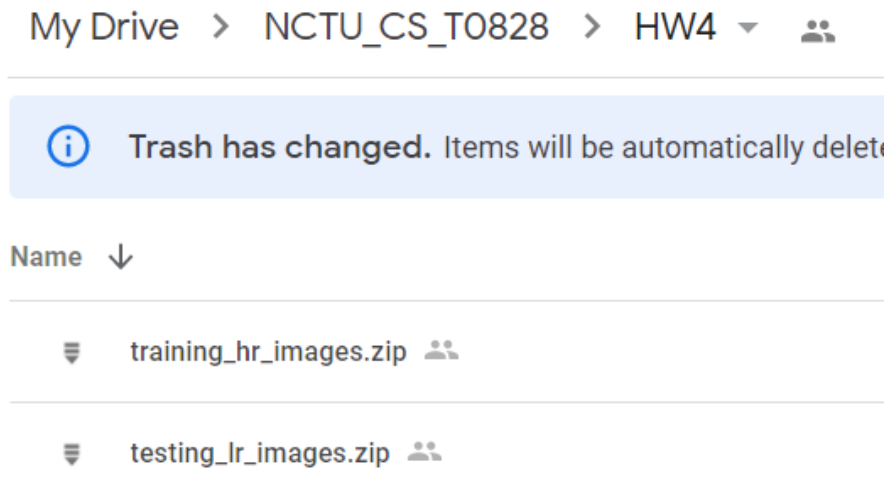


low-resolution  
image

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**



Ground Truth  
(PSNR, SSIM)



A+ [22]  
(22.92, 0.7379)



RFL [18]  
(22.90, 0.7332)



SelfEx [11]  
(23.00, 0.7439)



SRCNN [5]  
(23.15, 0.7487)

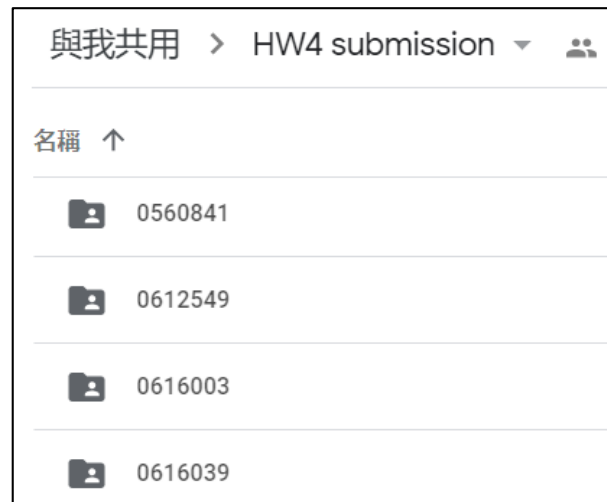
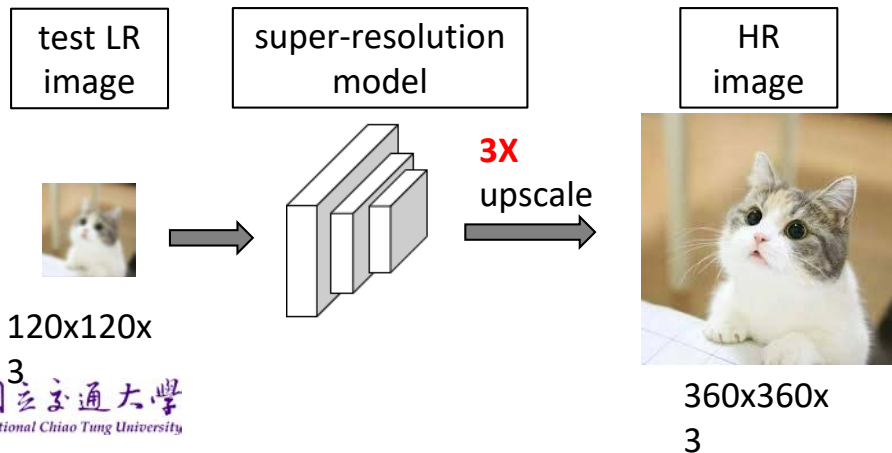


VDSR (Ours)  
(23.50, 0.7777)



# 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



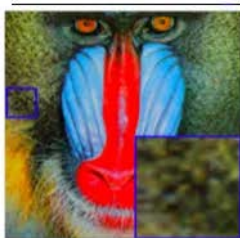


# Grading policy: Model performance (70 points)

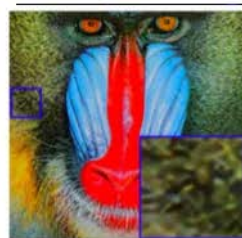
- Get at least 56% ( $70\% \times 0.8$ ) by scoring over the baseline



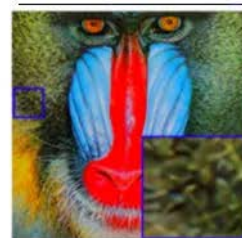
(a) Baboon Original



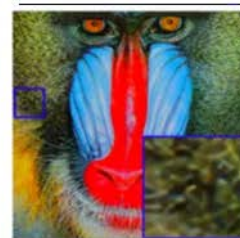
(b) Bicubic / 23.21db



(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

# Recommended

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

# Not recommended

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

Python

# Recommended

```
my_list = [1, 2, 3]
```

# Not recommended

```
my_list = [ 1, 2, 3, ]
```

Python

```
x = 5
```

```
y = 6
```

# Recommended

```
print(x, y)
```

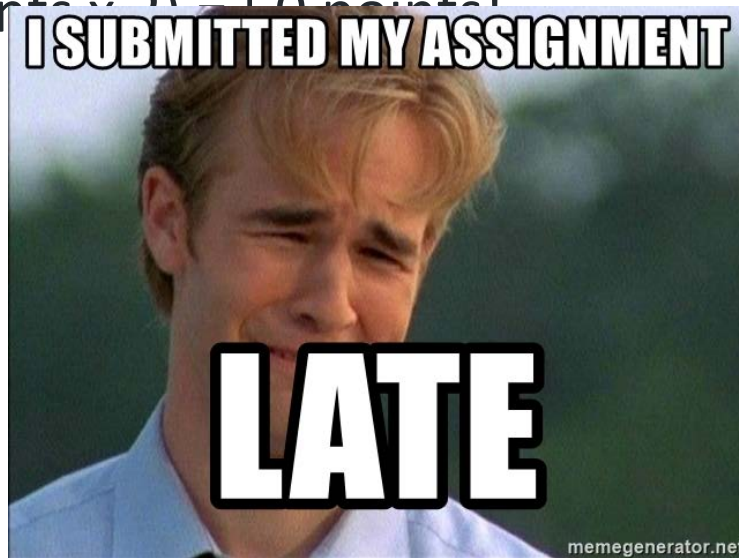
# Not recommended

```
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-  $(20 \text{ points} \times 2) = 50 \text{ points}$



# 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](mailto:lin@cs.nctu.edu.tw)
  - Jimmy: [d08922002@ntu.edu.tw](mailto:d08922002@ntu.edu.tw) (3pm-4pm, Thur., EC118)
  - 佳諭: [mylifeai1116@gmail.com](mailto:mylifeai1116@gmail.com) (3pm-4pm, Thur., EC118)
  - 玉霖: [oscar861201@gmail.com](mailto: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 University  
[d08922002@ntu.edu.tw](mailto: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
<b>Bronze</b>	Top 40%	Top 40%	Top 100	Top 10%
<b>Silver</b>	Top 20%	Top 20%	Top 50	Top 5%
<b>Gold</b>	Top 10%	Top 10	Top 10 + 0.2%*	Top 10 + 0.2%*



# Competitions on Kaggle






1. [APTOS 2019 Blindness Detection](#): Image classification
  2. [Handwritten Grapheme Classification](#): Image classification
  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

[Overview](#) [Data](#) [Notebooks](#) [Discussion](#) [Leaderboard](#) [Rules](#) [New Topic](#)

37		<b>Green is all u(may) need!!</b> Bibek 1y ago	Last comment 3d ago by felix_indra	13
2		<b>Public and Private Test Set</b> Quan 1y ago	Last comment 1mo ago by HieuDQM	2
12		<b>Are the labels of these testing images available?</b> estelle1728 1y ago	Last comment 3mo ago by Juan Widyaya	8
1		<b>why cv2.imread() run out of memory?</b> daniel 1y ago	Last comment 1mo ago by Philippe Modard	6
6		<b>112th place solution (Fastai student)</b> Hao He 1y ago	Last comment 3mo ago by Tan Pei Seng	3




# 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)




The banner features a background image of a man and a woman using an ophthalmoscope. The man is on the left, looking into the device, and the woman is on the right, looking through the eyepiece. The background is a solid green color.

 Featured Code Competition

**APTOS 2019 Blindness Detection**

Detect diabetic retinopathy to stop blindness before it's too late

 Asia Pacific Tele-Ophthalmology Society (APTOS) · 2,931 teams · a year ago

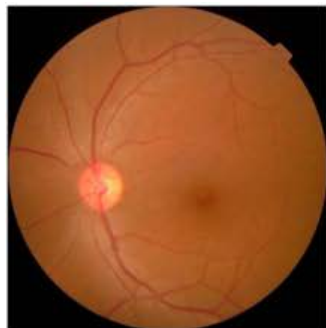
**\$50,000**  
Prize Money



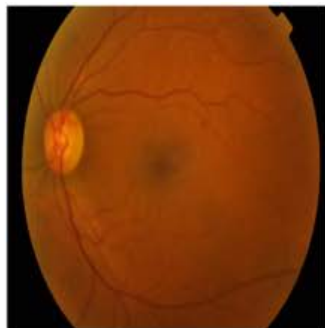


# 1. APTOS 2019 Blindness Detection

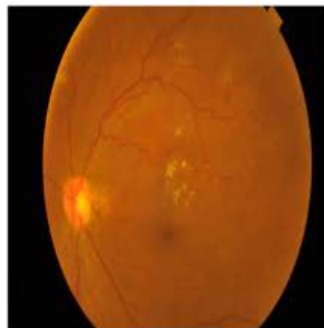
- Classify the **ratings** correctly for all images



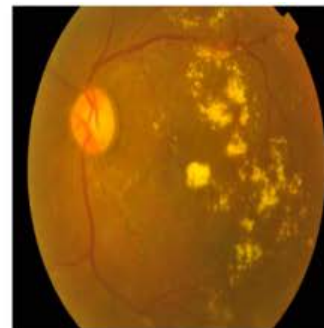
Normal



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)



Research Code Competition

### Bengali.AI Handwritten Grapheme Classification

\$10,000

Classify the components of handwritten Bengali

Prize Money



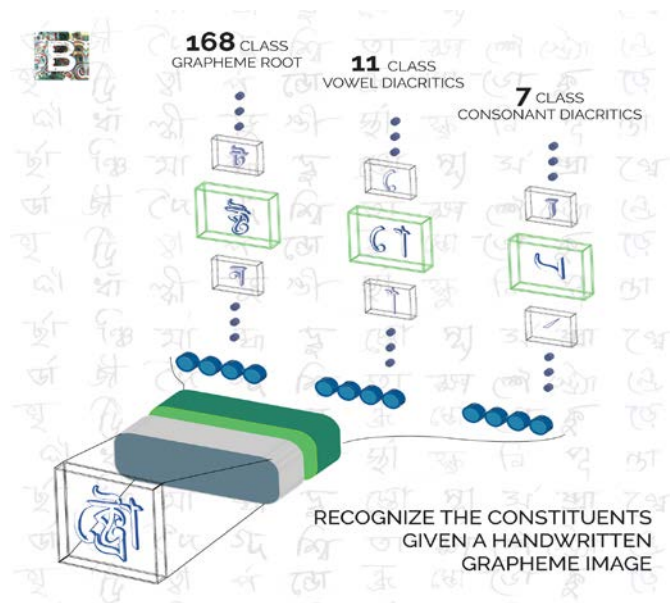
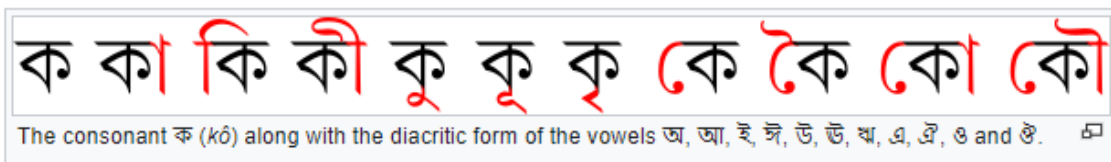
Bengali.AI · 2,059 teams · 8 months ago



國立交通大學  
National Chiao Tung University

## 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](#)



Research Code Competition

## Global Wheat Detection

Can you help identify wheat heads using image analysis?

\$15,000

Prize Money



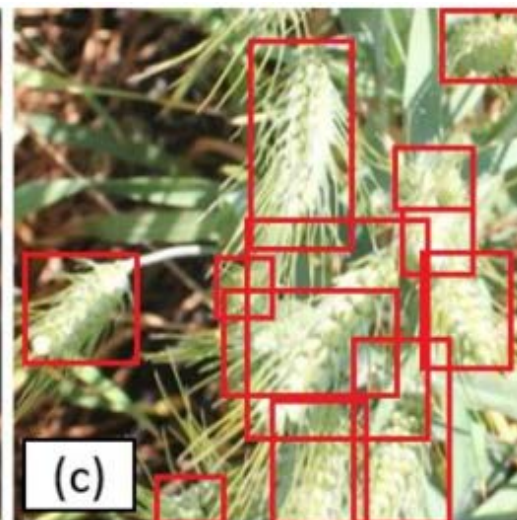
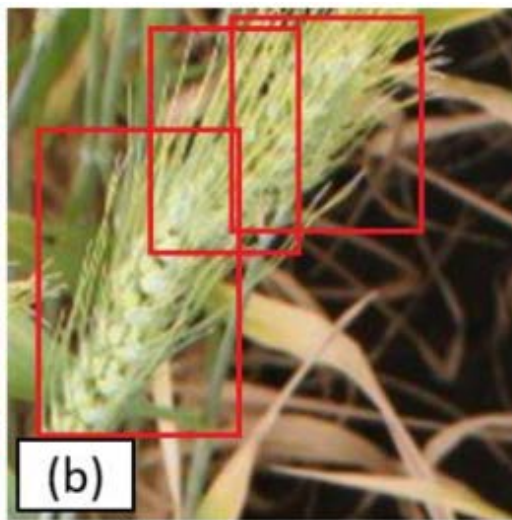
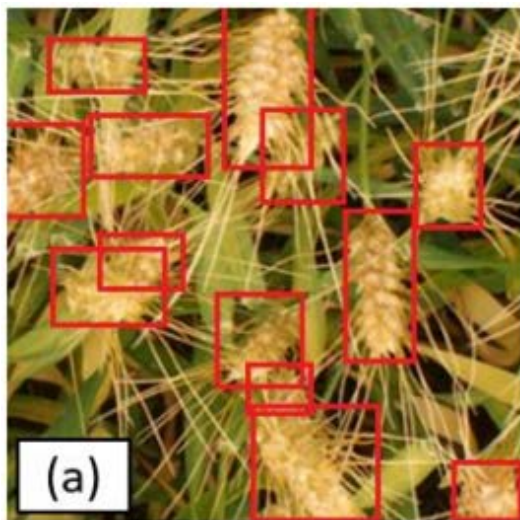
University of Saskatchewan · 2,245 teams · 3 months ago



國立交通大學  
National Chiao Tung University

### 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)



Featured Code Competition

### Severstal: Steel Defect Detection

Can you detect and classify defects in steel?

\$120,000

Prize Money



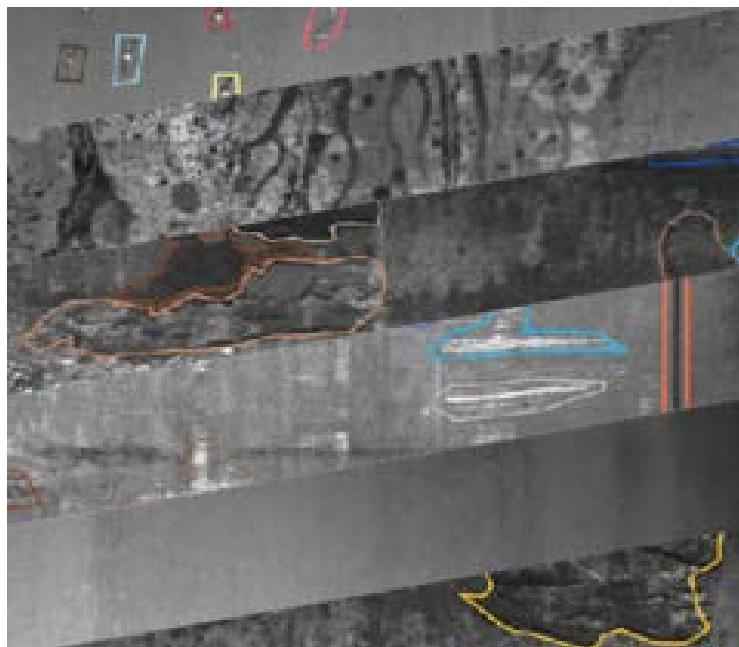
Severstal · 2,431 teams · a year ago



國立交通大學  
National Chiao Tung University

## 4. Severstal: Steel Defect Detection

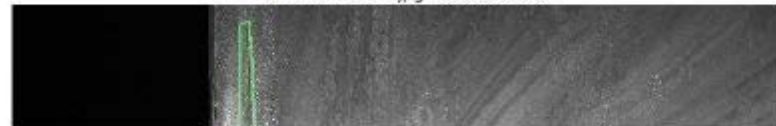
- Segment the **4 types of defect** correctly for all images



Train 0002cc93b.jpg has defect 1



Train 00031f466.jpg has defect 2



Train 000418bfc.jpg has defect 1 2





Train 000789191.jpg has defect 3 4





# Grading policy

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

162...	Gaurav Gooner Roy		0.56459	1	22d
162...	jimmy15923		0.56459	1	now

**Your First Entry** ⬆

Welcome to the leaderboard!

Your score represents your submission's accuracy. For example, a score of 0.7 in this competition indicates you predicted Titanic survival correctly for 70% of people.

**What next? You've got a few options:**

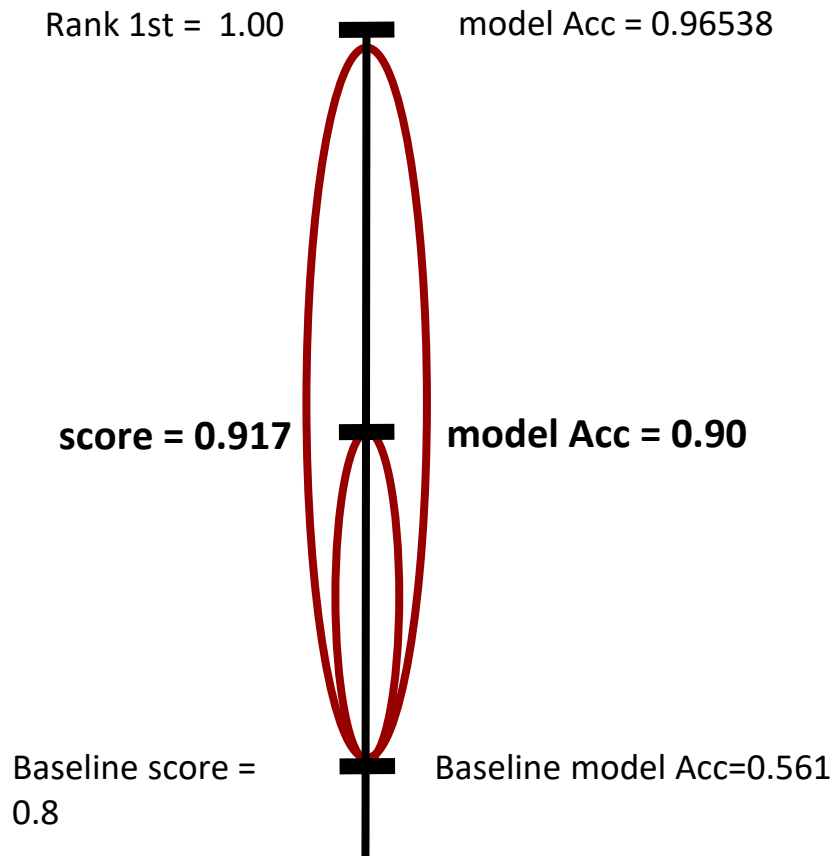
- 🤖 Learn skills that can improve your score in [our Intro to Machine Learning course by Dan Becker](#).
- 🔍 Check out [the discussion forum](#) to find lots of tutorials and insights from other competitors.
- 🏆 Find a new challenge by entering one of our [open, active competitions](#) or searching our [public datasets](#).





# 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 pros 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

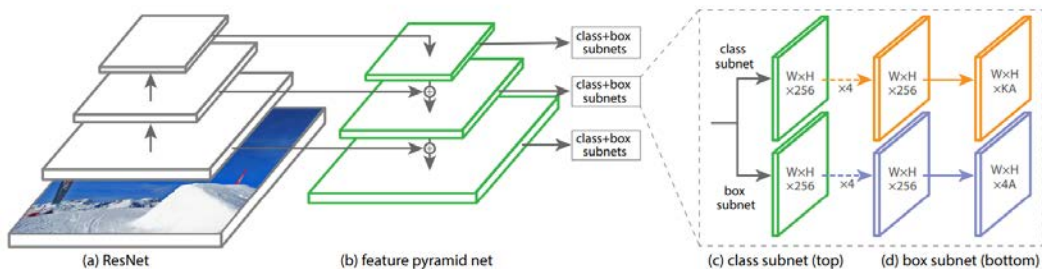
**Two-stage Detectors:** The dominant object detection is based on a two-stage neered in the Selective Search work [35

**One-stage Detectors:** OverFeat [30] was a modern one-stage object detector based on a convolutional neural network. 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



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



# 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 <sub>50</sub>	AP <sub>75</sub>	AP <sub>S</sub>	AP <sub>M</sub>	AP <sub>L</sub>
<i>Two-stage methods</i>							
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	<b>52.1</b>
<i>One-stage methods</i>							
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
<b>RetinaNet</b> (ours)	ResNet-101-FPN	39.1	59.1	42.3	21.8	42.7	50.2
<b>RetinaNet</b> (ours)	ResNeXt-101-FPN	<b>40.8</b>	<b>61.1</b>	<b>44.1</b>	<b>24.1</b>	<b>44.2</b>	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 ( <a href="#">google form</a> )	<b>12/5 Sat</b>
Select competition ( <a href="#">google form</a> )	<b>12/10 Thr</b>
Final presentation I & report submission	<b>12/24 Thr</b>
Final presentation II & report submission	<b>12/31 Thr</b>
Deadline of HW4	<b>01/07 Thr</b>



**Good luck!**

