



Institute of Electronics  
National Yang Ming Chiao Tung  
University  
Hsinchu, Taiwan

2024 ADAR Lab AI Training

# Final Project

## Image Super-Resolution



Architecture  
Development &  
Algorithm  
Refinement for  
Intelligent Computing

Presenter:

Lyu-Ming Ho

Advisor:

Juinn-Dar Huang, Ph.D.

August 22nd, 2024

# Outline

---

- Intro of Super Resolution
- Performance Evaluation: PSNR
- Task Description
- Deadline & Grading & Requirement
- Submission
- Report
- Hint
- Appendix

# Intro of Super Resolution

- A low level computer vision task
- LR → HR (no unique solution , 人眼覺得好看就是好)
- Restore the sharp edges and textures from LR images
- No unique ground truth in real life, infinite solution space

## Image Super Resolution



- Ex: 240 X 240 → 1080 X 1080
- Ex: HD → UHD

# Performance Evaluation: PSNR

- PSNR (Peak Signal to Noise Ratio)
- 通常PSNR 高於30 dB可以視為還不錯的標準
- 這部分評分的code我會提供(demo.py)，請不要擔心
- 如果有需要，請參考：

[https://blog.csdn.net/weixin\\_43624538/article/details/95201186](https://blog.csdn.net/weixin_43624538/article/details/95201186)

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2$$

$$PSNR = 10 \cdot \log_{10} \left( \frac{MAX_I^2}{MSE} \right)$$

# Task Description

---

- 做出一個模型可以將input圖片長寬各放大**3倍**，而且model需要可以吃下**任何的input size**，因為Set5每張照片大小都不一樣，只要跑得過Set5就好
- Example:
  - 假設input為640 x 360，output需為1920 x 1080
  - 假設input為1280 x 720，output需為3840 x 2160
- Training dataset (不限):
  - DIV2k, Flickr2k(連結appendix(4)，請自己載)，也可以不要用這兩個dataset，如果需要自己切 training image size或自己切validation set都是可以的(換言之，你想對dataset做什麼都可以，你也可以自己找其他的dataset 來train，只要model符合spec就好，能在demo環境中跑起來就可以)
- Testing dataset:
  - Set5(會提供)

# Deadline & Grading

---

- Deadline:
  - 9/2, 開學第一天。
- Grading:
  - Set5 PSNR (**50%**) , 跟你們寄給我的信裡面寫的數值需一樣 (appendix(3))
    - ›  $20 \leq \text{PSNR} < 29$  (12.5 %)
    - ›  $29 \leq \text{PSNR} < 31$  (25 %)
    - ›  $31 \leq \text{PSNR} < 32$  (37.5 %)
    - ›  $32 \leq \text{PSNR}$  (50 %)
  - Report (**35 %**)
  - Special case (**15 %**)
    - › Model 處理一張 $256 \times 256 \rightarrow 768 \times 768$  的FLOPS算力  $< 11.1$  G FLOPs，且 Set5 PSNR  $> 31.5$  (RGB space) , code我也會先幫各位寫好，大家只要執行我提供的 demo.py 就可以知道自己的model有沒有符合規定

# Rules

---

- 基本上跟midterm project一樣，但是要注意的是，你們不可以拿網路上別人的model跟pretrained weight就拿來交。
- 你也不能把Set5拿來train，因為我會拿其他testing dataset來看一下你的model表現到底合不合理
- 因為基本上網路上很多model我都跑過以及看過了，我看code就知道這是哪個model，所以要抄的話不要抄太像，重點是要有自己的巧思。

# Submission

---

- Submit your code to: [holyuming.ee12@nycu.edu.tw](mailto:holyuming.ee12@nycu.edu.tw)
  - Email Subject: [重要] <名字> 繳交AI training final project , 看appendix (3)
  - <your English name>\_final\_project.zip
- <your English name>\_final\_project
  - model.pth 請參照appendix (1), 檔名就是這樣不要改
  - model.py 請參照appendix (1), 檔名就是這樣不要改
  - requirements.txt (可免) 請參照appendix (1), 檔名就是這樣不要改
  - <中文名字>\_final\_report.pdf
- If you submit the wrong format or disobey submission rules:
  - 被扣10分
  - 然後你專題會被當掉 (開玩笑的) 😂 😂
  - 如果格式錯了我會很麻煩，拜託各位了 😱 😱

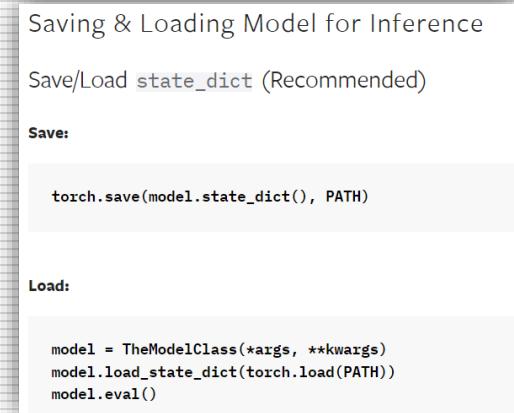
# Report

---

- **畫圖 (20 %)**
  - 畫你的model architecture圖，如: 看appendix (2)
- **說明你做了什麼 (不限) (25 %)**
  - 你用了什麼額外的training data
  - 你對training dataset做了什麼augmentation，切成很多256 x 256去訓練，水平flip ...
  - 你看了什麼SR論文或什麼csdn博主文章，阿感想是什麼
  - 你覺得怎樣做比較好
  - ...，想寫什麼都可以，上述只是舉例，不一定要，重點是你做了什麼
- **對整個ai training 課程的感想 (5 %)**
  - 說什麼都可以，幫助明年的教學
  - 有寫有分，正面或負面評價都可以，發自內心就好，不影響分數
  - 最少50字
  - 認真寫，真的對我們有幫助，再加分0 ~ 5 %
- **你在此次的final project學到了什麼 or 觀察到什麼 (10 %)**
  - 寫什麼都可以，不用太多，這個會認真看，讓我感覺你都沒學到或寫不好會不給分 😊

# Note (重要)

- 檔名請不要改，就叫**model.py & model.pth**
- **model.py** 不能更動**model class name**
- 如果你們需要下載額外的package or library，請寫在**requirements.txt**裡面
- 你們**training**的結束之後請用下圖方式來儲存模型:其中**PATH = ‘model.pth’**，也就是**torch.save(model.state\_dict(), ‘model.pth’)**



Save model when training:

```
torch.save(model.state_dict(), 'model.pth')
```

demo.py

```
model = SuperResolution().to(device)

state_dict = torch.load('model.pth', map_location=device)
model.load_state_dict(state_dict, strict=False)
model.eval()
```

# Hint

- 可以多問問題，真的遇到問題就問，不要害怕問問題，雖然我很忙，但我有看到我還是會回，阿如果我也不會的那你就得靠自己了
- 可以自己試試看把dataset的圖片切小一些，再送進model training。
- 如果不想要算力太高，在pixelshuffle之後不要加太多東西。
- Model深度在SR這個Task扮演一個很crucial的角色，但誠如你們所學，太深的model會怎樣？會有梯度消失的問題，所以residual link對training很重要。但也不建議拉太多
- 注重算力的同學可以注意，普遍來看，convolution 比 transformer算力多很多，大家可以注意自己模型兩者的比例
- nn.PixelShuffle: <https://pytorch.org/docs/stable/generated/torch.nn.PixelShuffle.html>
- Paper with code: <https://paperswithcode.com>
- 我覺得不錯的且有名的model提供給大家學習模仿他的想法：
  - SRCNN, FSRCNN, RCAN, SwinIR, HAT, EDSR, DLRN ...

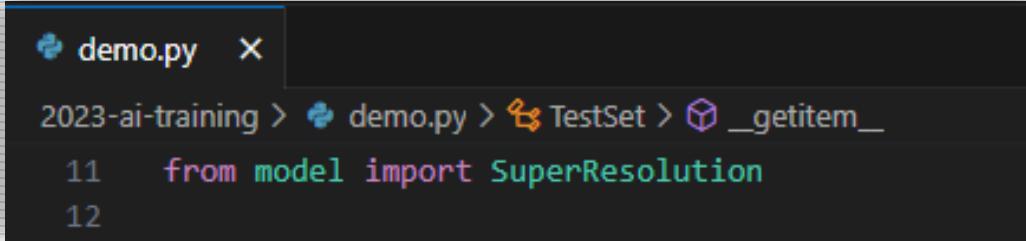
# Appendix (1)

- model.py，model的名稱必須為 SuperResolution，demo的時候我會用 from model import SuperResolution 來import你的model，如圖：

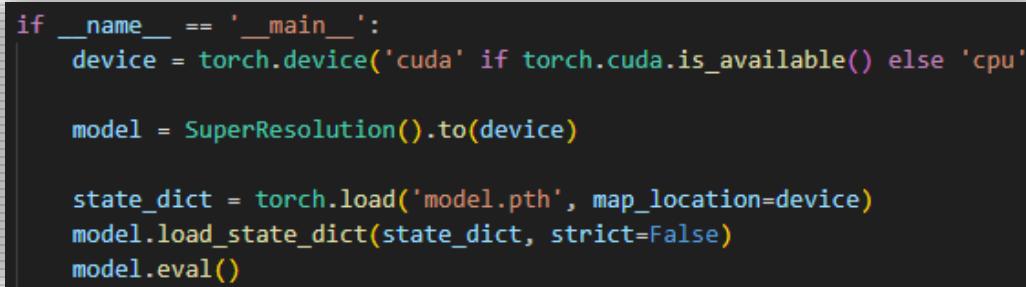
```
model.py  x
2023-ai-training > model.py > SuperResolution > forward
 1  import torch
 2  import torch.nn as nn
 3  from thop import profile
 4
 5
 6  class SuperResolution(nn.Module):
 7      def __init__(self):
 8          super(SuperResolution, self).__init__()
 9          self.conv1 = nn.Conv2d(3, 64, 3, 1, 1)
10          self.conv2 = nn.Conv2d(64, 32, 3, 1, 1)
11          self.conv3 = nn.Conv2d(32, 27, 3, 1, 1)
12          self.PS = nn.PixelShuffle(3)
13          self.relu = nn.ReLU(inplace=True)
14          self.conv4 = nn.Conv2d(3, 3, 3, 1, 1)
15
16      def forward(self, x):
17          x = self.relu(self.conv1(x))
18          x = self.relu(self.conv2(x))
19          x = self.relu(self.conv3(x))
20          x = self.PS(x)
21          return x
22
```

# Appendix (1)

- demo.py，此檔案我會提供，請不要動此檔案(read-only)，如果動了，會造成我這邊demo出來的結果與你那邊不符合。
- 在demo.py當中，我會用這個方式去呼叫你的model，建議大家去看一下demo.py如何執行。



```
demo.py  x
2023-ai-training > demo.py > TestSet > _getitem_
11     from model import SuperResolution
12
```



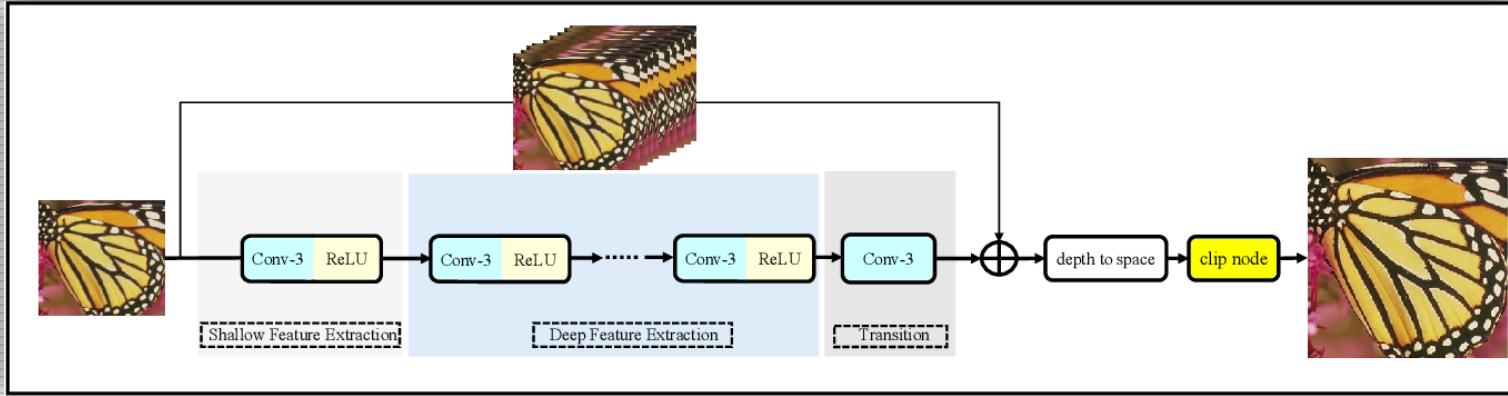
```
if __name__ == '__main__':
    device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')

    model = SuperResolution().to(device)

    state_dict = torch.load('model.pth', map_location=device)
    model.load_state_dict(state_dict, strict=False)
    model.eval()
```

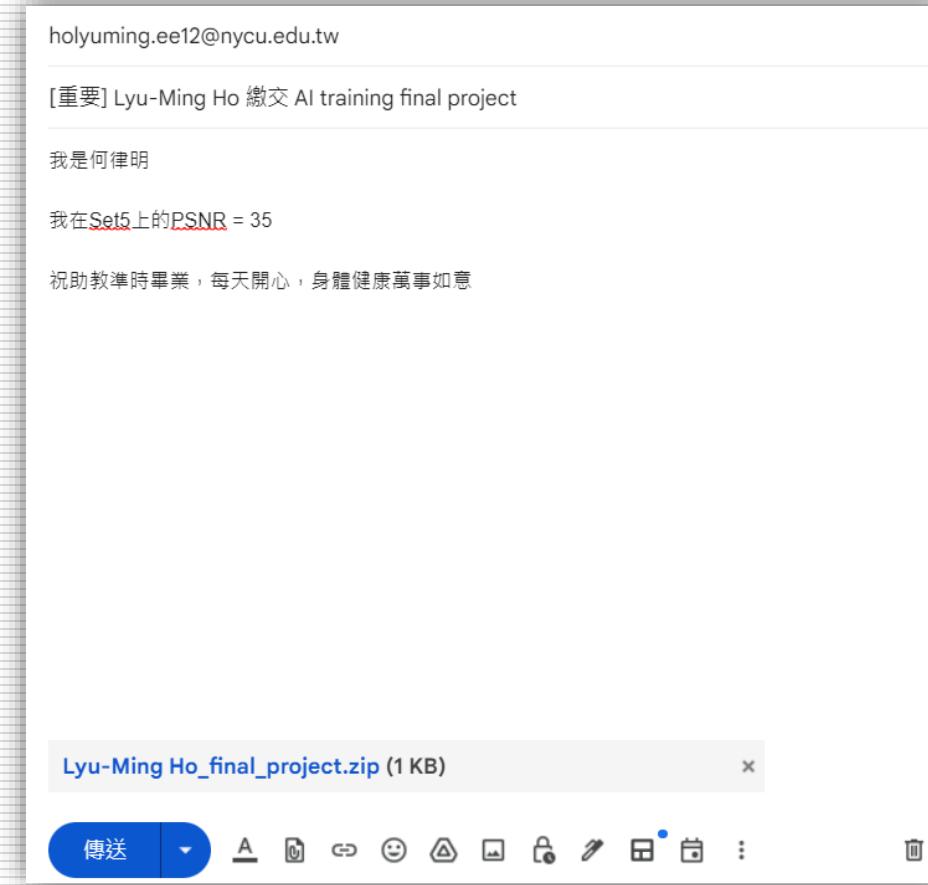
# Appendix (2)

- Report model架構圖，至少要讓我看的出來你做了什麼module block，當然也可以標一下tensor 的shape，大概畫一下，不需要太華麗沒關係，如：



# Appendix (3)

- 寄信模板如下：請記得標註你們Set5的PSNR值



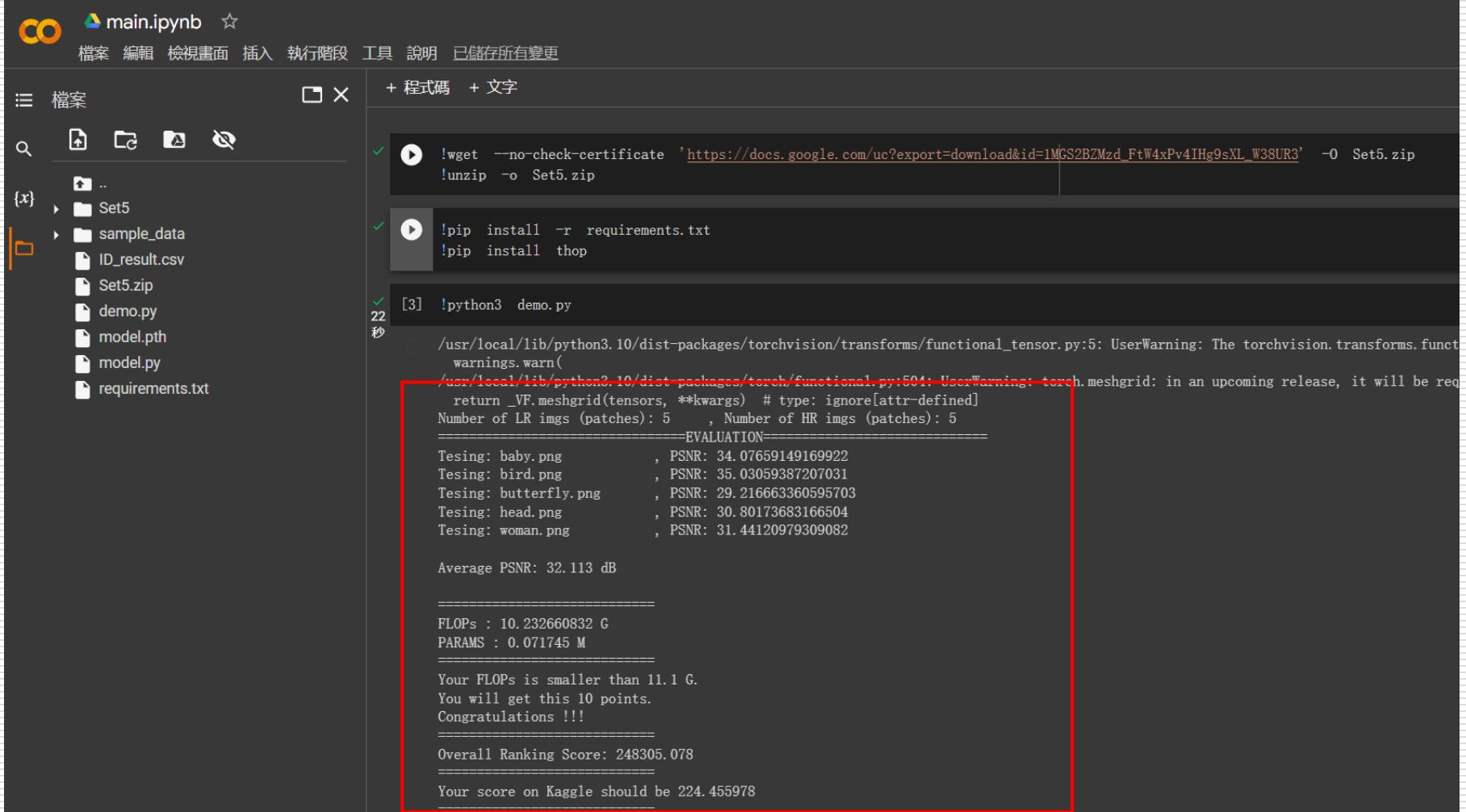
# Appendix (4)

---

- DIV2k: <https://data.vision.ee.ethz.ch/cvl/DIV2K/>
- Flickr2k: <https://cv.snu.ac.kr/research/EDSR/Flickr2K.tar>

# Appendix (5), main.ipynb

將main.ipynb放到colab上，再將model.pth, model.py, requirements.txt拉到左邊的檔案區，直接執行main.ipynb就可以了。



```
!wget --no-check-certificate 'https://docs.google.com/uc?export=download&id=1MGS2BZMzd_FtW4xPv4IHg9sXL_W38UR3' -O Set5.zip
!unzip -o Set5.zip

!pip install -r requirements.txt
!pip install thop

[3] !python3 demo.py
22 秒
/usr/local/lib/python3.10/dist-packages/torchvision/transforms/functional_tensor.py:5: UserWarning: The torchvision.transforms.functional module has moved to the main torchvision package. Please update your imports. This will happen on 2020-08-14. To suppress this warning, you can use warnings.warn(
/warnings.warn(
/usr/local/lib/python3.10/dist-packages/torch/functional.py:504: UserWarning: torch.meshgrid: in an upcoming release, it will be renamed to _VF.meshgrid(tensors, **kwargs) # type: ignore[attr-defined]
Number of LR imgs (patches): 5 , Number of HR imgs (patches): 5
=====EVALUATION=====
Tesiing: baby.png , PSNR: 34.07659149169922
Tesiing: bird.png , PSNR: 35.03059387207031
Tesiing: butterfly.png , PSNR: 29.216663360595703
Tesiing: head.png , PSNR: 30.80173683166504
Tesiing: woman.png , PSNR: 31.44120979309082

Average PSNR: 32.113 dB

=====
FLOPs : 10.232660832 G
PARAMS : 0.071745 M
=====

Your FLOPs is smaller than 11.1 G.
You will get this 10 points.
Congratulations !!!
=====

Overall Ranking Score: 248305.078
=====

Your score on Kaggle should be 224.455978
```

---

Thank you