# 影像處理、電腦視覺及深度學習概論 (Introduction to Image Processing, Computer Vision and Deep Learning)

#### Homework 2

TA: 張議隆

Gmail: nckubot65904@gmail.com

Office Hour: 14:00~16:00, Mon.

10:00~12:00, Fri.

At CSIE 9F Robotics Lab.

## Notice (1/2)

- Copying homework is strictly prohibited!! Penalty: Both individuals will receive a score of 0!!
- Due date => 09:00:00, 2023/12/19 (Tue.)

Do not submit late, or the following points will be deducted:

- ➤ Submit within seven days after the deadline, and your score will be reduced by half.
- ➤ If you submit after this period, you will receive a score of 0.
- You must attend the demonstration, otherwise your score will be 0. The demonstration schedule will be announced on NCKU Moodle.
- You must create GUI, otherwise your point will be deducted.
- Upload to => 140.116.154.28 -> Upload/Homework/Hw2
  - ➤ User ID: opencvd12023 Password: RL2023opencv
- Format
  - > Filename: Hw2\_StudentID\_Name\_Version.rar
    - Ex: Hw2\_F71234567\_林小明\_V1.rar
    - If you want to update your file, you should update your version to be V2,
    - Ex: Hw2\_F71234567\_林小明\_V2.rar
  - Content: Project folder \*( Excluding the pictures )
    - \*Note: Remove your "Debug" folder to reduce file size.

## Notice (2/2)

- Python (recommended):
  - > Python 3.8 (<a href="https://www.python.org/downloads/">https://www.python.org/downloads/</a>)
  - **→** Opency-contrib-python (3.4.2.17)
  - ➤ Matplotlib 3.7.3
  - ➤ UI framework: pyqt5 (5.15.10)
  - > Pytorch 2.1.0
  - > Torchvision 0.16.0
  - ➤ Torchsummary 1.5.1
  - > Tensorboard 2.14.0
  - > Pillow 10.1.0

## Assignment scoring (Total: 100%)

- 1. (20%) Hough Circle Transform (出題: You)
  - 1.1 (15%) Draw Contour
  - 1.2 (5%) Count Rings

\* Don't fix your image path
(There is another dataset for demonstration)
Load image 請用下面Function 來讀取路徑
QFileDialog.getOpenFileName

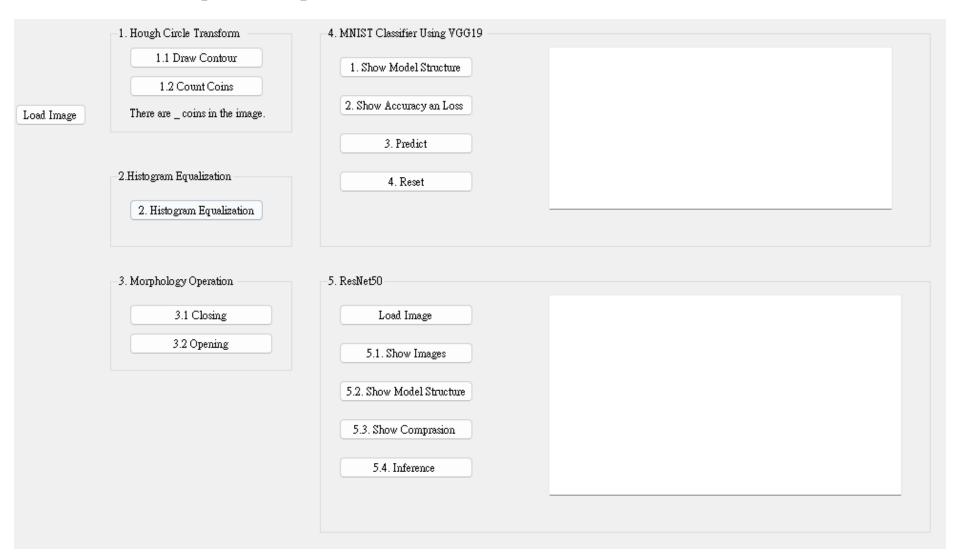
獲取打開的檔路徑

- 2. (20%) Histogram Equalization (出題: Eric)
- 3. (20%) Morphology Operation (出題: Hsiang)
  - 3.1 (10%) Closing
  - 3.2 (10%) Opening
- 4. (20%) Training a MNIST Classifier Using VGG19 with BN (出題: Shang)
  - 4.1 (6%) Load Model and Show Model Structure.
  - 4.2 (6%) Show Training/Validating Accuracy and Loss.
  - 4.3 (8%) Use the Model with Highest Validation Accuracy to Run Inference, Show the Predicted Distribution and Class Label.
- 5. (20%) Train a Cat-Dog Classifier Using ResNet50 (出題: Shan)
  - 5.1 (5%) Load the dataset and resize images
  - 5.2 (5%) Show the structure of ResNet50 model
  - 5.3 (5%) Improve ResNet50 with Random-Erasing and Compare the accuracies of 2 ResNet50 models on validation dataset
  - 5.4 (5%) Use the trained model to run inference and show the predicted class label

Load Image

# Assignment scoring (Total: 100%)

• Use one UI to present 5 questions.



# 1. (20%) Hough Circle Transform

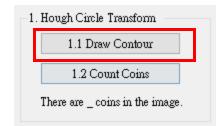
(出題:You)

- 1.1 (15%) Draw Contour
- 1.2 (5%) Count Rings

## 1.1 Hough Circle Transform – Draw Contour

(出題:You)

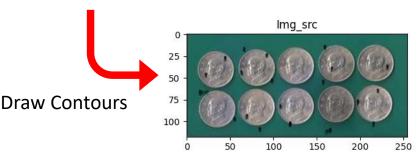
- ☐ Given: One color image, "coins.jpg"
- Q: 1) Draw Contour: Using OpenCV functions to find the contours of coins in image.
  - 2) **Circle Center**: Show center points of the circle with most votes.

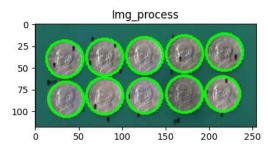


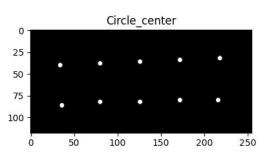
- ☐ Hint: Textbook Chapter 6, p.153 ~ p.161
  - 1. RGB → Grayscale
  - 2. Remember to remove the noise. (Use cv2.GaussianBlur(5, 5))
  - 3. Using circle detection function to get result. (Use cv2.HoughCircles)
  - 4. Display the original image, processed image, and circle center image at the same time.



coins.jpg: 255x118





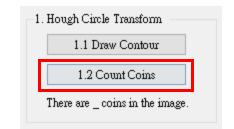


## **1.2 Hough Circle Transform – Count Coins**

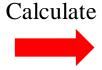
☐ Given: One color image, "coins.jpg"

Q: 3) Count coins: Using OpenCV functions to find how many coins in the image.

☐ Hint: Textbook Chapter 6, p.153 ~ p.161 Calculate how many coins and show on your UI. (出題:You)







There are 10 coins in the image.

coins.jpg: 255x118

## 2. (20%) Histogram Equalization

- ➤ Given: A single grayscale image will be provided as input.
  - Q: Perform histogram equalization on a grayscale image to enhance its contrast using OpenCV and manual(PDF&CDF).

(出題: Eric)

Part 1: Using OpenCV

- 1) Load histoEqual.png (click "Load Image" button).
  - Hint: your\_image\_path = **filedialog.askopenfilename**()
- 2) Perform histogram equalization using cv2.equalizeHist() function from OpenCV to process image.
- 3) The equalized image will be displayed alongside the original image for comparison in the upper quadrants.
- 4) The histograms of both the original and the equalized images will be shown in the lower quadrants.
  - Hint: using **plt.bar()** to display the frequency (y-axis) of each grayscale value (x-axis).

#### Part 2: Using PDF and CDF (tutorial)

- 1) The **Probability Density Function (PDF)** represents the frequency of each grayscale level in the image.
- 2) The Cumulative Distribution Function (CDF) is the cumulative sum of the PDF and is used to map the old grayscale values to new ones.
- 3) The program calculates the histogram of the grayscale image using numpy.histogram().
- 4) The PDF is obtained from the normalized histogram.
- 5) The CDF is calculated by cumulatively summing the PDF.
- 6) A lookup table is created based on the rounded CDF values, it is applied to the original image to create a new equalized image.
- 7) The histogram of the manually equalized image is also plotted, showing the redistributed grayscale frequencies.
- 8) Click "2. Histogram Equalization" button to show result.

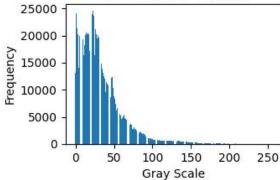
## 2. (20%) Histogram Equalization

> Result:

Original Image



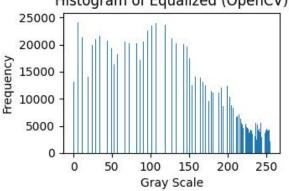
Histogram of Original



Equalized with OpenCV



Histogram of Equalized (OpenCV)

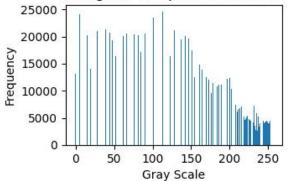


**Equalized Manually** 

(出題: Eric)



Histogram of Equalized (Manual)



➤ GUI:

Load Image

Histogram Equalization

2. Histogram Equalization

# 3. Morphology Operation (20%)

(出題:Hsiang)

- 3.1 (10%) Closing
- 3.2 (10%) Opening

## 3. Morphology Operation (1/3)

#### (出題:Hsiang)

## **Steps of Erosion Operation:**

### 1. Define Structuring Element

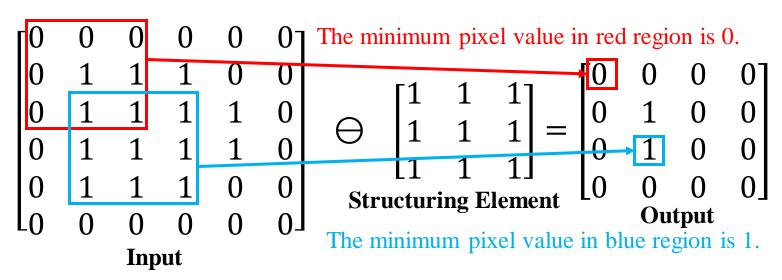
Choose a structuring element, typically a small square or rectangular convolution kernel. The size and shape of this element determine the effect of the erosion.

#### 2. Scanning

Initiate the process by sliding the structuring element from the top-left corner of the image. Sequentially cover each pixel in the image, moving both horizontally from left to right and vertically from top to bottom.

#### 3. Pixel Update

For each covered region, update the pixel values covered by the structuring element to the minimum pixel value in that region.



# 3. Morphology Operation (2/3)

#### (出題:Hsiang)

## **Steps of Dilation Operation:**

### 1. Define Structuring Element

Choose a structuring element, typically a small square or rectangular convolution kernel. The size and shape of this element determine the effect of the erosion.

#### 2. Scanning

Initiate the process by sliding the structuring element from the top-left corner of the image. Sequentially cover each pixel in the image, moving both horizontally from left to right and vertically from top to bottom.

#### 3. Pixel Update

For each covered region, update the pixel values covered by the structuring element to the maximum pixel value in that region.

0	0	0	0	0	ر0	The maximum pixel value in red region is 1.
0	1	1	1	0	0	$[1 \ 1 \ 1]$ $[1 \ 1 \ 1]$
0	1	1	1	1	0	$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \end{vmatrix}$
0	1	1	1	1	0	$\begin{array}{ c c c c c c c c c c c c c c c c c c c$
0	1	1	1	0	0	Structuring Element \[ \begin{array}{c ccccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
<b>L</b> 0	0	0	0	0	07	Output
Input					The maximum pixel value in blue region is 1.	

## 3. Morphology Operation (3/3)

### 1. Closing Operation:

- An operation performed by first applying dilation followed by erosion:
   Closing(A) = Erosion(Dilation(A))
- The purpose is to fill small holes in an image while preserving the shape and size of large holes and objects in the image.

#### 2. Opening Operation:

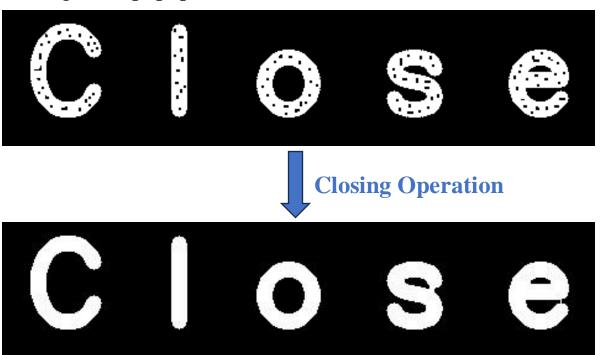
- An operation performed by first applying erosion followed by dilation :
   Opening(A) = Dilation (Erosion(A))
- The purpose is to remove small objects and thin lines from an image while preserving the shape and size of larger objects in the image

(出題: Hsiang)

## 3.1 Closing (10%)

(出題:Hsiang)

- 1. Given: "closing.png"
- 2. Constraint: Can not use OpenCV Function cv2.dilate(), cv2.erosion(), cv2.morphologyEx()
- 3. Question: Perform opening operation on the image
  - 1) Using "Load image" button to load RGB image.
  - 2) Convert the RGB image to grayscale.
  - 3) Binarize the grayscale image, assigning values of 0 or 255 only. (threshold = 127)
  - 4) Pad the image with zeros based on the kernel size (K=3).
  - 5) Perform the dilation operation using a 3x3 all-ones structuring element.
  - 6) Perform the erosion operation using a 3x3 all-ones structuring element.
  - 7) Show the image in a popup window.



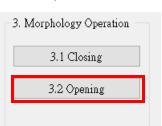


## 3.1 Opening (10%)

(出題:Hsiang)

- 1. Given: "opening.png"
- 2. Constraint: Can not use OpenCV Function cv2.dilate(), cv2.erosion(), cv2.morphologyEx()
- 3. Question: Perform opening operation on the image
  - 1) Using "Load image" button to load RGB image.
  - 2) Convert the RGB image to grayscale.
  - 3) Binarize the grayscale image, assigning values of 0 or 255 only. (threshold = 127)
  - 4) Pad the image with zeros based on the kernel size (K=3).
  - 5) Perform the erosion operation using a 3x3 all-ones structuring element.
  - 6) Perform the dilation operation using a 3x3 all-ones structuring element.
  - 7) Show the image in a popup window.





## 4. Training a MNIST Classifier Using VGG19 with BN (20%)

- 4.1 Load Model and Show Model Structure. (6%)
- 4.2 Show Training/Validating Accuracy and Loss. (6%)
- 4.3 Use the Model with Highest Validation Accuracy to Run Inference, Show the Predicted Distribution and Class Label. (8%)

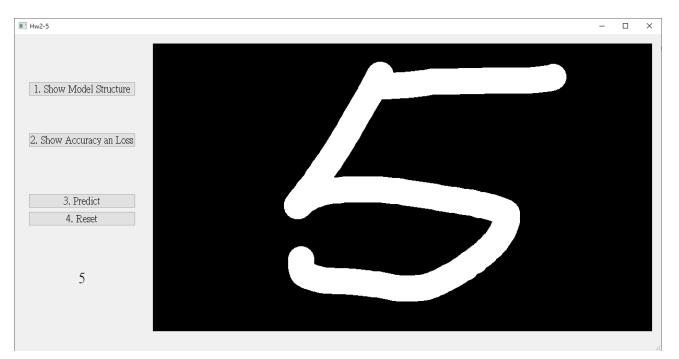


Figure: GUI example

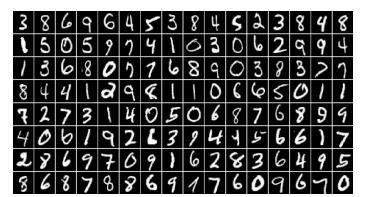
(出題:Shang)

## 4.0 Training a MNIST Classifier Using VGG19 with BN (20%)

- Requirement (出題: Shang)
  - 1) Train VGG19 model with batch normalization (BN) using PyTorch.
  - 2) Download dataset using torchvision.datasets.MNIST() (tutorial)
    - Training data: 60000 images
    - Validation data: 10000 images
    - Resize image to (32, 32)
  - 3) Parameters
    - At least 30 epochs.
    - Cross entropy loss
    - Adam optimizer
  - 4) Record training/validation loss and accuracy in .jpg or .png format.
  - 5) In the submitted file, you need to include
    - Weight file for VGG19 with BN in .pth format. (File size is approximately 540MB)
    - Figure of training/validating loss and accuracy in .jpg or .png format.
    - Code for your GUI program
    - Code for model training.
  - 6) Please do not include image data in the submitted file.

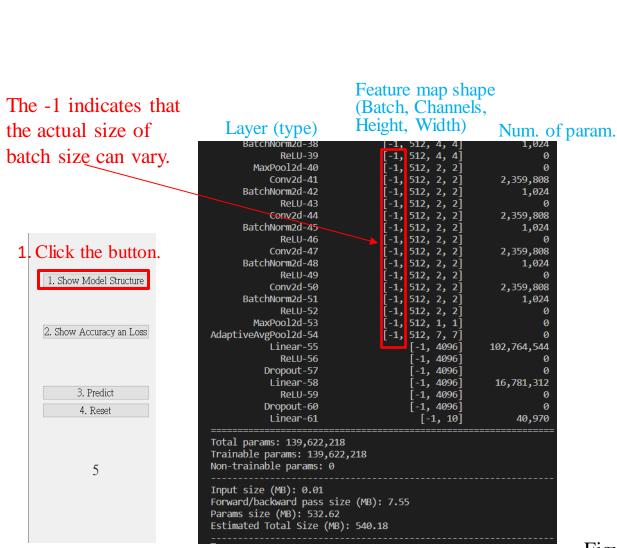
#### R. Reference

- 1) VGG19
- 2) Batch Normalization



## 4.1 Show the Structure of VGG19 with BN (6%)

- 1. Click the button "1. Show Model Structure"
- 2. Show the VGG19 with BN model on terminal using torchsummary.summary().



(出題:Shang) Input Image (32x32x3) 32x32x64 + BN + ReLU 32x32x64 + BN + ReLU pool-1: 16x16x64 16x16x128 + BN + ReLU 16x16x128 + BN + ReLU pool-2: 8x8x128 8x8x256 + BN + ReLU All convolution pool-3: 4x4x256 4x4x512 + BN + ReLU filter size is 3x3 4x4x512 + BN + ReLU 4x4x512 + BN + ReLU 4x4x512 + BN + ReLU pool-4: 2x2x512 2x2x512 + BN + ReLU Flatten Here pool-5: 1x1x512 Adaptive pool: 7x7x512 FC: 4096 + ReLU + Dropout Input Layer FC: 4096 ReLU + Convolution + ReLU Dropout Max-pooling Adaptive-pooling FC: 10 Fully connected(FC) + ReLU Softmax Output + sigmoid

Figure: the Structure of VGG19 with BN

## 4.2 Show Training/Validating Accuracy and Loss (6%)

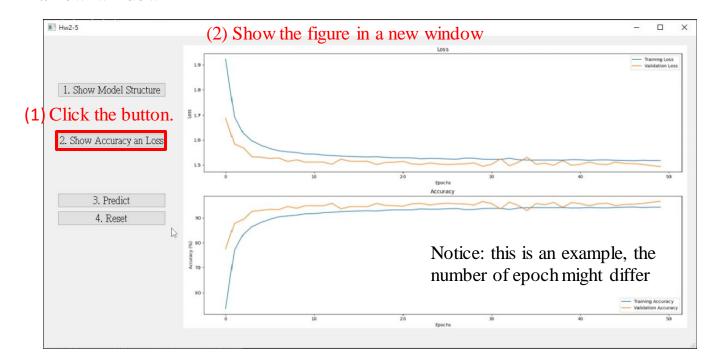
#### 1. At home:

(出題:Shang)

- 1) Download the training and validation datasets. (tutorial)
- 2) Training and validating VGG19 with BN at least 30 epochs at home (<u>tutorial</u>) and record the training/validating accuracy and loss in each epoch (<u>tutorial</u>).
- 3) If your validation accuracy is low, you can try
  - Adjust the learning rate of the optimizer.
  - Change the data augmentation techniques used.
- 4) Save weight file with highest validation accuracy.
- 5) Use <u>matplotlib.pyplot.plot()</u> to create a line chart for the <u>training</u> and <u>validating loss</u> and accuracy values and save the figure.

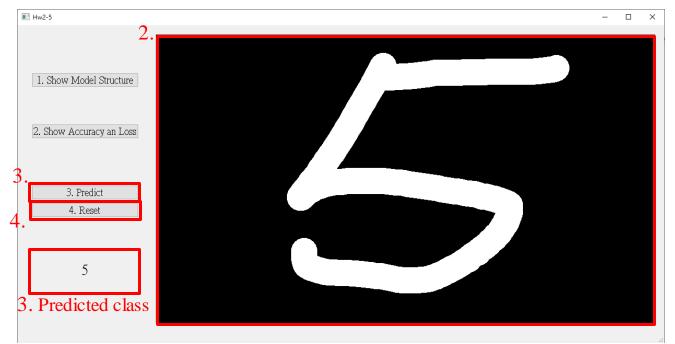
#### 2. When the demo:

- (1) Click the button "2. Show Accuracy and Loss"
- (2) Show the saved figure of Training/Validating loss and accuracy in a new window

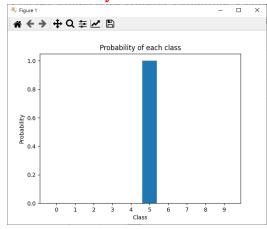


# 4.3 Use the Model with Highest Validation Accuracy to Run Inference, Show the Predicted Distribution and Class Label. (8%) (出題: Shang)

- 1. Load the model with highest validation accuracy which trained at home.
- 2. Draw a number on graffiti board using mouse. (tutorial)
  - Background: black
  - Pen: white
- 3. Click the button "3. Predict" to run inference on the image you drew.
  - Show the predicted class label on the GUI.
  - Show the probability distribution of model predictions using a histogram in a new window.
- 4. Click the button "4. reset" to clear the graffiti board.

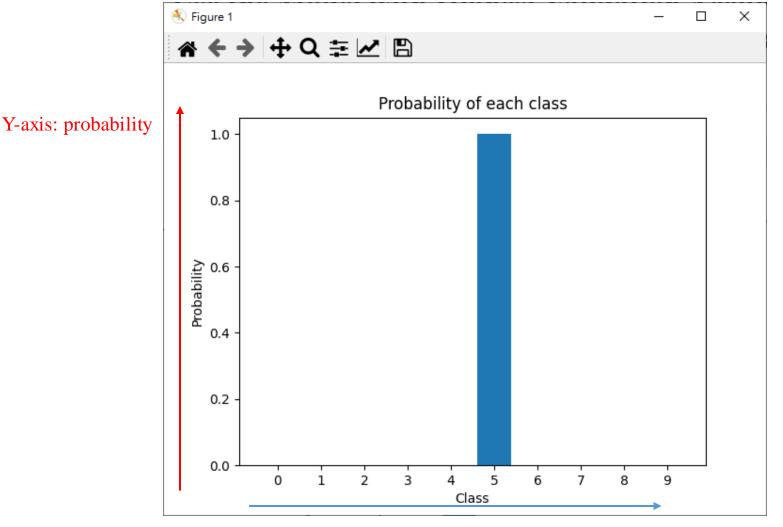


#### 3. Probability of each class



# 4.3 Use the Model with Highest Validation Accuracy to Run Inference, Show the Predicted Distribution and Class Label. (6%) (出題: Shang)

The probability distribution of model prediction using a histogram.

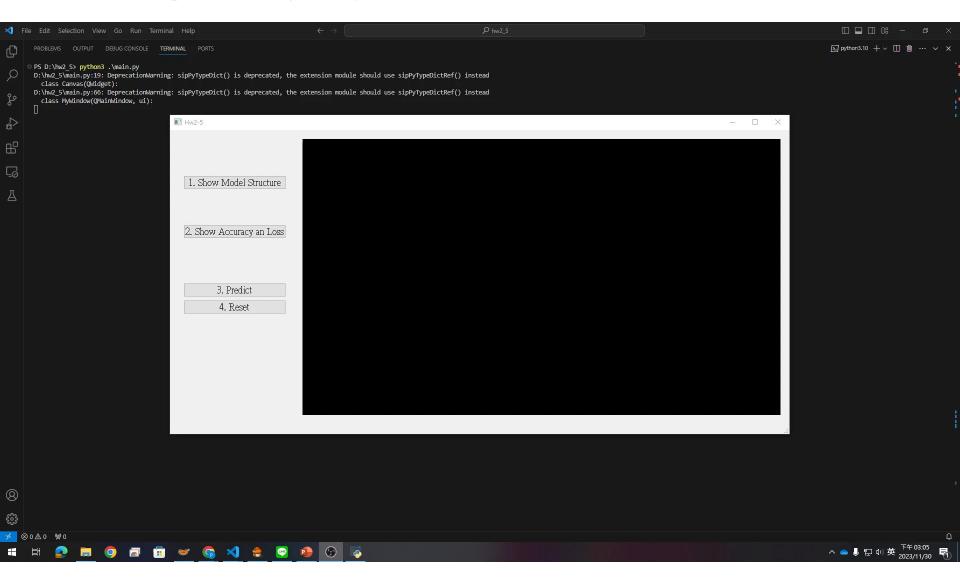


X-axis: class name

## 4. Training a MNIST Classifier Using VGG19 – Example Video

• This is an example illustrating the objectives from  $4.1 \sim 4.3$ .

(出題:Shang)



## 5. Train a Cat-Dog Classifier Using ResNet50 (20%) (出題: Shan)

- 5.1 (5%) Load the dataset and resize images
- 5.2 (5%) Show the structure of ResNet50 model
- 5.3 (5%) Improve ResNet50 with Random-Erasing and Compare the accuracies of 2 ResNet50 models on validation dataset
- 5.4 (5%) Use the trained model to run inference and show the predicted class label

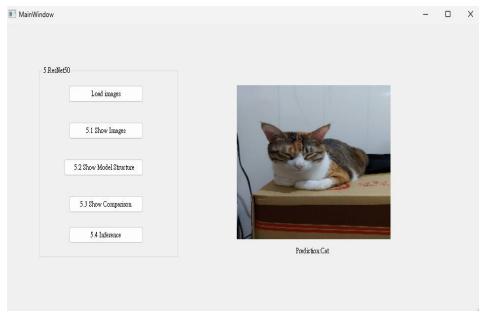


Figure: GUI Example

## 5.0 Train a Cat-Dog Classifier Using ResNet50

(出題: Shan)

- 1. Objective
  - 1) Learn how to train a ResNet50 model to classify images of cats and dogs using PyTorch (tutorial)
- 2. Download Cats and Dogs Dataset from FTP
  - 1) Data type: JPG images
  - 2) 2 classes: Cat and Dog
  - 3) Datasets
    - (1) Training dataset: 16,200 JPG images in total.
    - (2) Validation dataset: 1,800 JPG images in total.
    - (3) Inference dataset: 10 JPG images in total.

      It is for testing the inference function in your GUI program.

#### 3. In the submitted file

your homework file.

1) Organize the files in this structure:

#### R. Reference

- 1) Deep Residual Learning for Image Recognition
- 2) Kaggle Cats and Dogs Dataset

## 5.1 (5%) Load the dataset and resize images

(出題: Shan)

- 1) At home:
  - (1) Load the inference dataset
    - → Hint:
      - (a) PyTorch (tutorial): torch.utils.data.Dataset
  - (2) Resize images to 224×224×3c (RGB)
    - → Hint:
      - (a) PyTorch (tutorial): torchvision.transform
  - (3) Click the button "1. Show Images"
  - (4) Get 1 image from each class in the inference dataset
  - (5) Show images in a new window
    - → Hint: use matplotlib.pyplot functions to show images (tutorial):
      - (a) figure()
      - (b) imshow()
      - (C) subplot()
      - (d) title()
- 2) When the demo:
  - (1) Click the button "1. Show Images"
  - (2) Show images in a new window



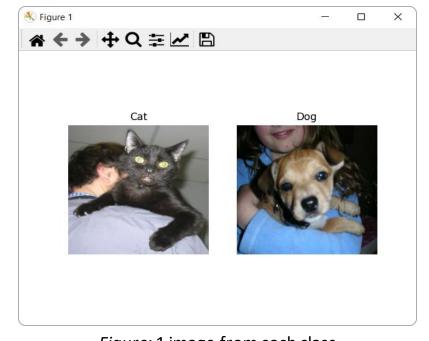


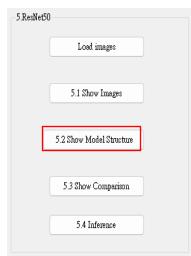
Figure: 1 image from each class

Notice: this is an example, the images might differ

## 5.2 (5%) Show the structure of ResNet50 model

(出題: Shan)

- 1) At home:
  - (1) Build a ResNet50 model
    - → Hint:
      - (a) PyTorch: torchvision.models.resnet50()
  - (2) Replace the output layer to a FC (Fully Connected) layer of 1 node with a Sigmoid activation function
    - → Hint:
      - (a) PyTorch ( $\underline{\text{tutorial}}$ ): torch.nn.Linear(2048, 1), torch.nn.Sigmoid If the class label of Cat is 1, the output value (range: 0 ~ 1) should be close to 1 for cat images, and vice versa.
  - (3) Run the function to show the structure in the terminal
    - → Hint:
      - (a) PyTorch: torchsummary
- 2) When the demo:
  - (1) Click the button "3. Show Model Structure"
  - (2) Run the function to show the structure in the terminal



## 5.2 (5%) Show the structure of ResNet50 model

(出題: Shan)



## 5.3 (5%) Improve ResNet50 with Random-Erasing

(出題: Shan)

- 1) At home: Set up **Random-Erasing** in codes for model training (train.py)
  - (1) Train 2 ResNet50 models with training dataset
    - → Hint:(a) PyTorch (tutorial): write a for loop to validate the model
      - (a) With Random-Erasing
      - (b) Without Random-Erasing

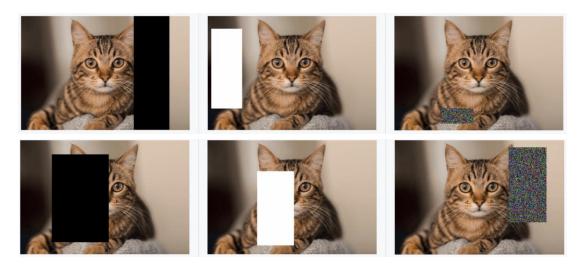


Figure 1: Examples of the use of Random-Erasing

2) When the demo: Show your codes about Random-Erasing in train.py

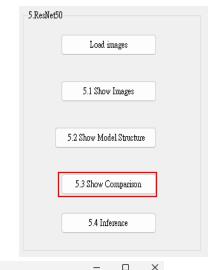
```
transform = transforms.Compose([
    transforms.Resize(224),
    transforms.CenterCrop(224),
    transforms.RandomHorizontalFlip(),
    transforms.RandomVerticalFlip(),
    transforms.ToTensor(),
    transforms.RandomErasing(),
])
```

R. Reference

5.3 (5%) Compare the accuracies of 2 ResNet50

models on validation dataset

- 1) At home:
  - (1) Validate 2 ResNet50 models with validation dataset
    - → Hint:
      - (a) PyTorch (tutorial): write a for loop to validate the model
  - (2) Plot the accuracy values with a bar chart
  - (3) Save the figure
- 2) When the demo:
  - (1) Click the button "4. Show Comparison"
  - (2) Show the saved figure of accuracy comparison in a new window



(出題: Shan)

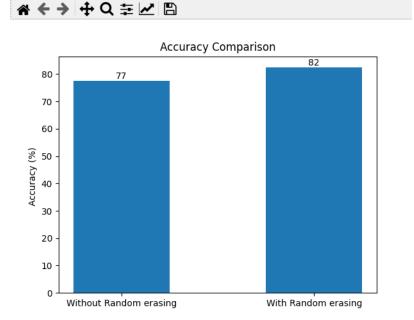


Figure 1: Accuracy Comparison

Notice: this is an example, the numbers might differ

# 5.4 (5%) Use the better-trained model to run inference and show the predicted class label

(出題: Shan)

- 1) At home:
  - (1) Load the trained model
    - → Hint:
      - (a) PyTorch: torch.nn.Module.load\_state\_dict()
  - (2) Click the button "Load Image" to select 1 image arbitrarily
    - → Hint: PyQt5.QtWidgets.QFileDialog.getOpenFileName()
  - (3) Show the loaded image in the GUI
  - (4) Resize the loaded image to 224×224×3c (RGB)
  - (5) Click the button "5. Inference" to run inference on the resized image
    - → Hint:
      - (a) PyTorch: pass an image when calling torch.nn.Module

object to run

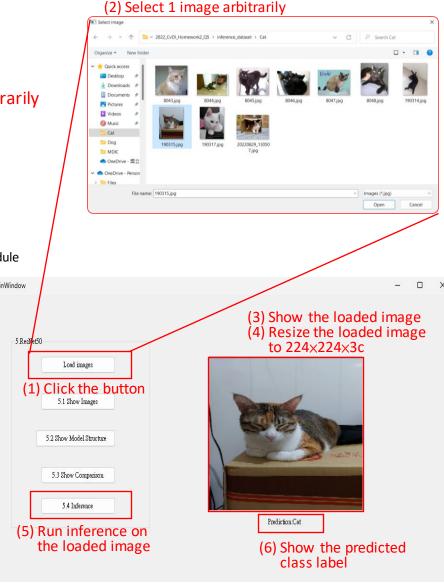
inference, ex: trained\_model(img)

- (6) Show the predicted class label
  - → Hint: decide the class label with a threshold of the output value.

Ex: class label = 
$$\begin{cases} Cat, output < thresh \\ Dog, output \ge thresh \end{cases}$$

$$thresh = 0.5$$

2) When the demo: repeat the process



## 5. Train a Cat-Dog Classifier Using ResNet50 - Demo Video

• This is an example illustrating the objectives from  $5.1 \sim 5...$ 

