

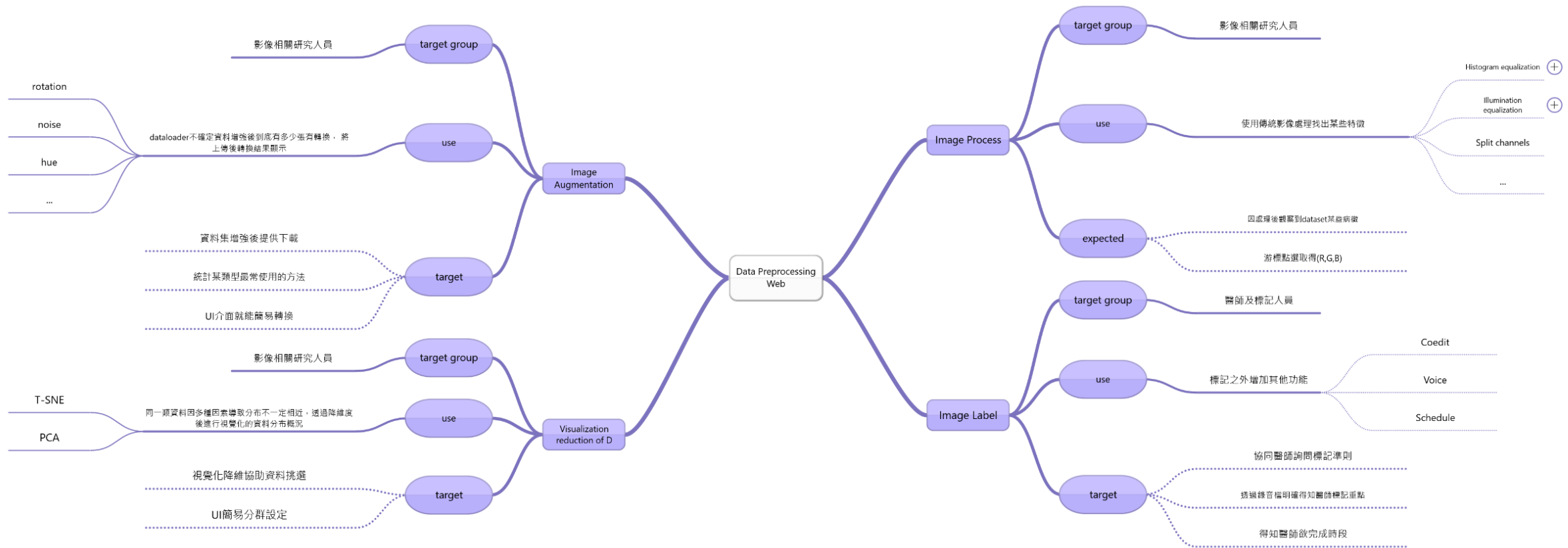
Outline

- Image augmentation
- Visualization reduction of dimensionality
- Image process
- Label image

Motivation

- ML和DL領域中需要”資料”來進行Model的訓練。
 - 這些資料”Image”為大宗。
- Image會受不同拍攝機器、光線、角度.....造成品質不一。
 - 進行特殊處理
 - 疊合不同處理
- 訓練資料random挑選的話，在某些情況下會需要依靠”運氣”來挑選。
- 節省與醫師討論資料標記之準則時間。

Map



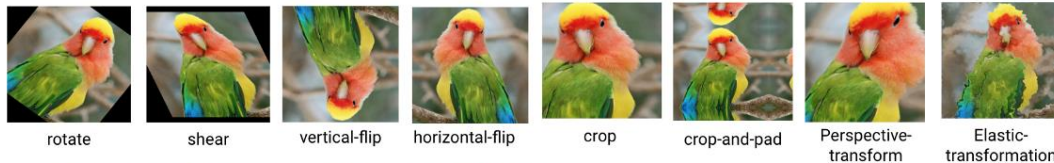
Target groups

- 目標客群:
 - 做影像相關領域人員—Ex：Lab成員
 - 協同資料標記-- Ex：醫師

Image augmentation

Base Augmentations

Geometry based



Color based



Noise / occlusion

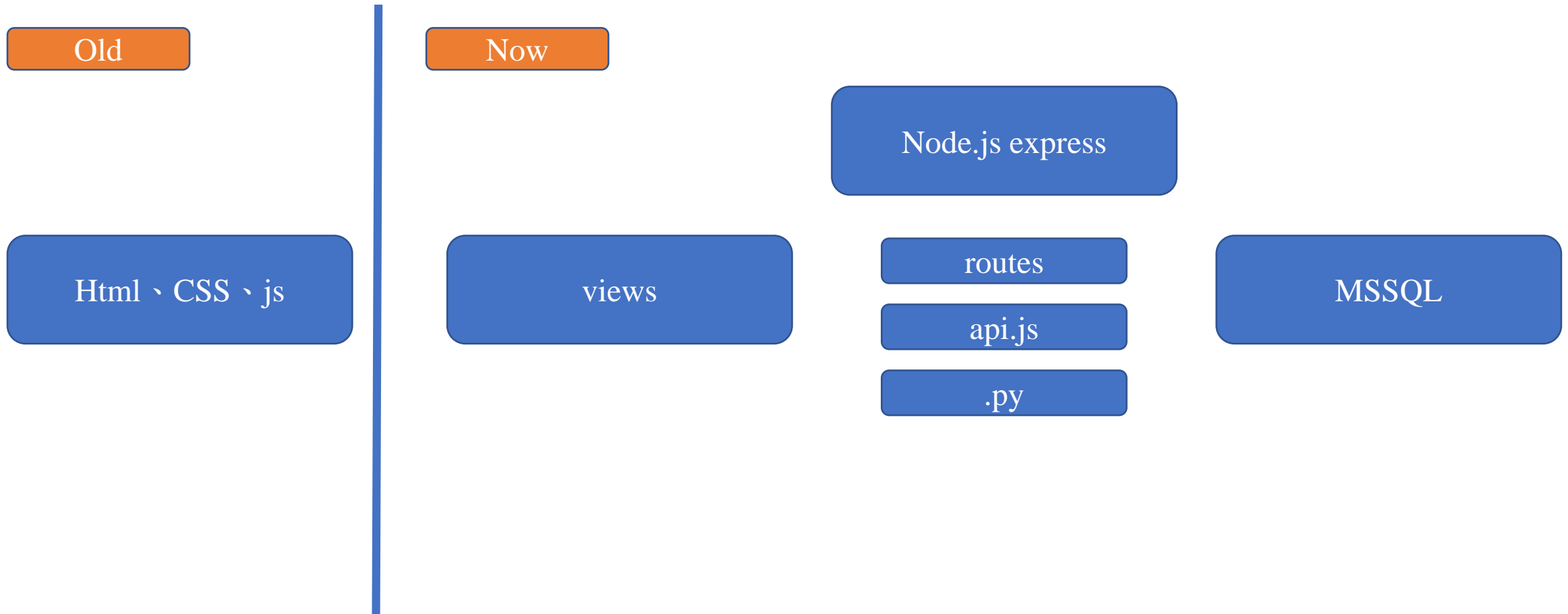


Weather



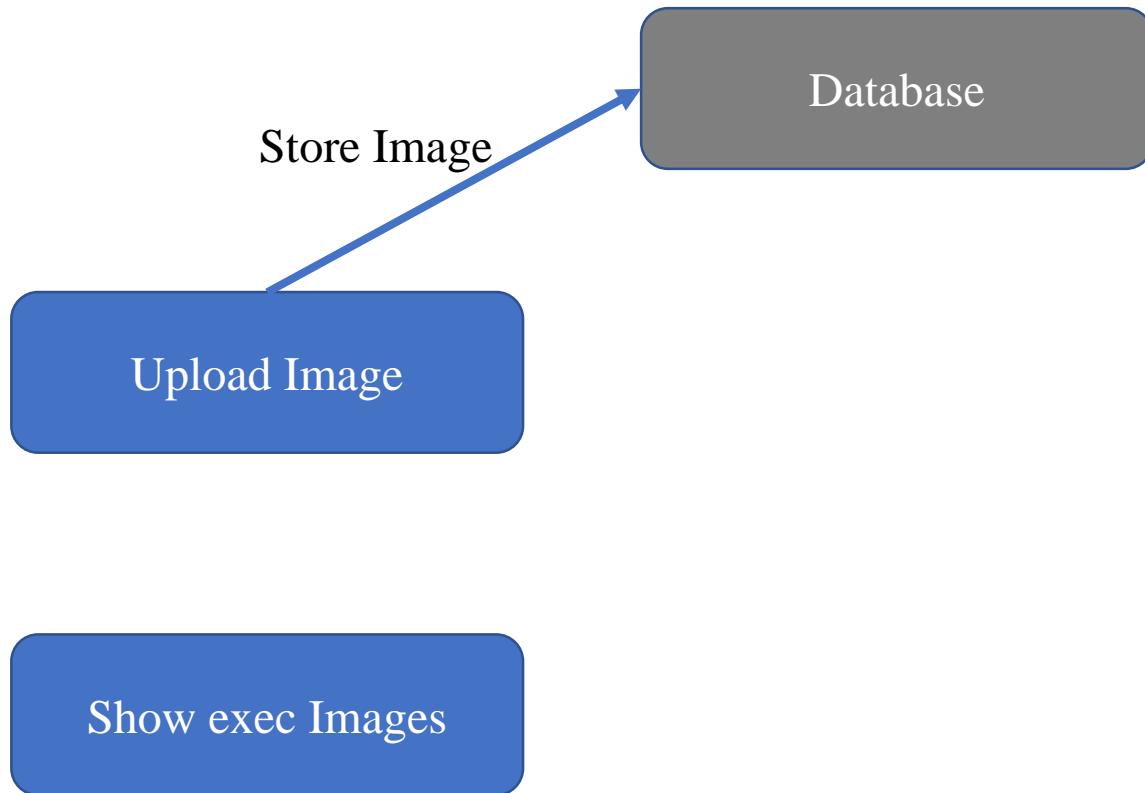
- 資料增強:
 - DataLoader是在原圖上隨機產生增強效果
 - 需另外轉換type查看轉換前後差異
 - 不確定多少張確實有轉換
- 預期:
 - 勾選調整轉換
 - 增強後資料下載

Html transfer to Node.js express(hbs)





Augmentation



上傳資料集




upload

Auto Augmentation


資料增強
1.上傳資料集

Upload multiple images

選擇檔案 4檔案



201.jpg 202.jpg 203.jpg

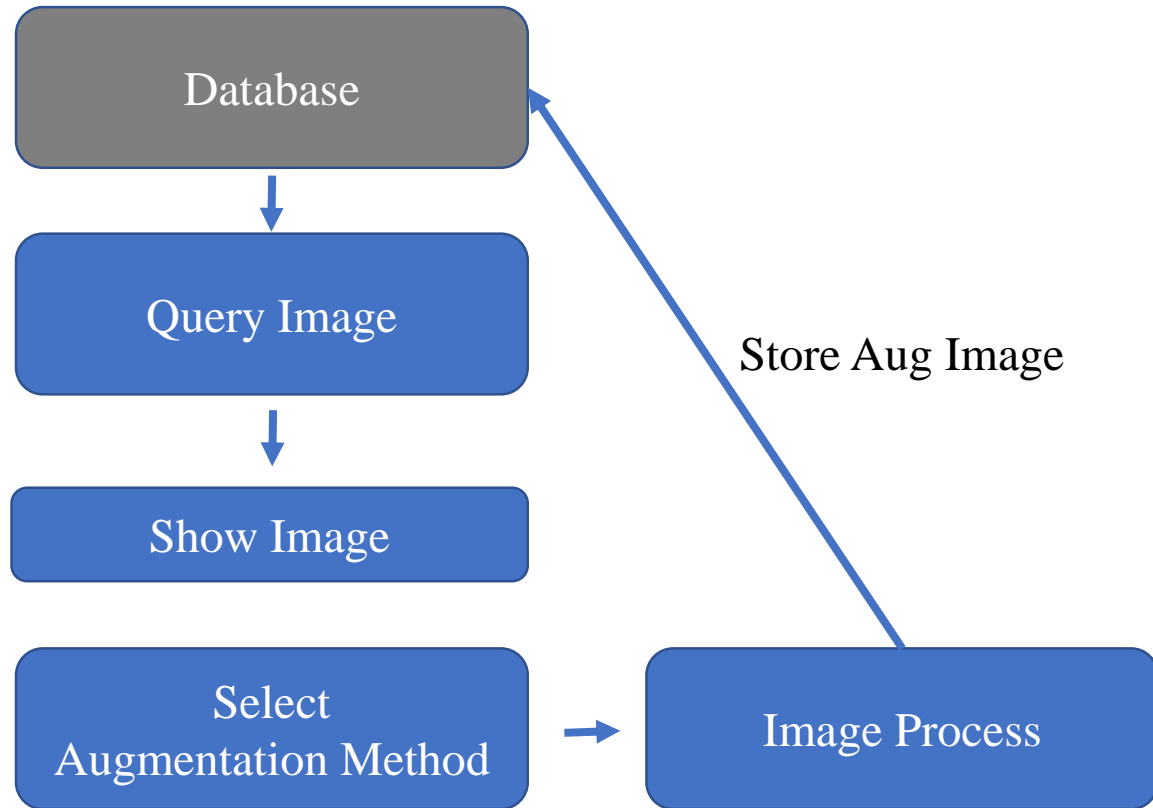


204.jpg

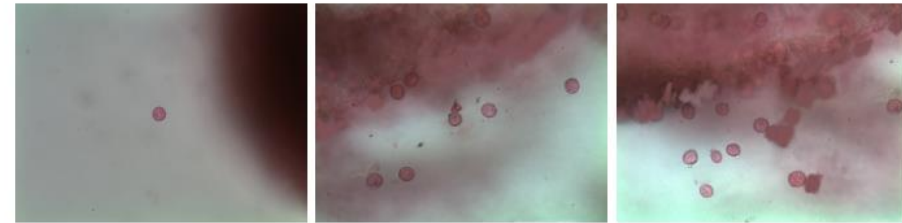
Upload

AIIA LAB 2022

Augmentation



資料増強
3.調整方法相關參數
4.組合多種方法
5.執行



0

1

2



3

選擇增強方式

☒ 垂直翻轉 ☐ 水平翻轉 ☒ 彩色直方圖均化

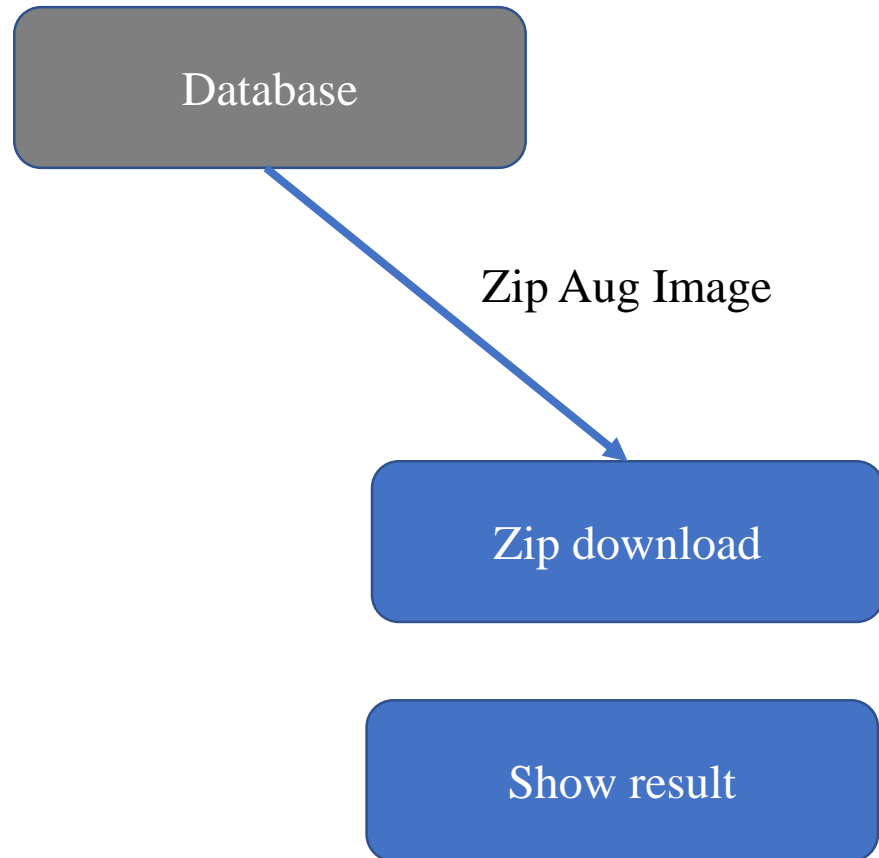
亮度調整 1.6

角度調整 7

submit

AIIA LAB 2022

Augmentation



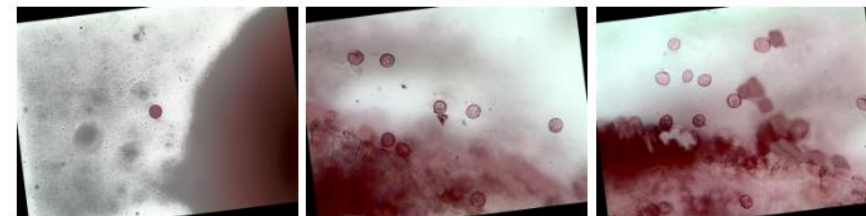
一鍵下載

downloads

Auto
Augmentation

資料增強
6. 一鍵下載

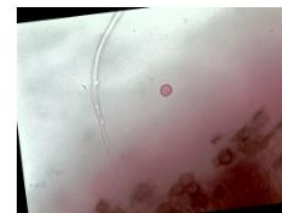
下載.ZIP



0

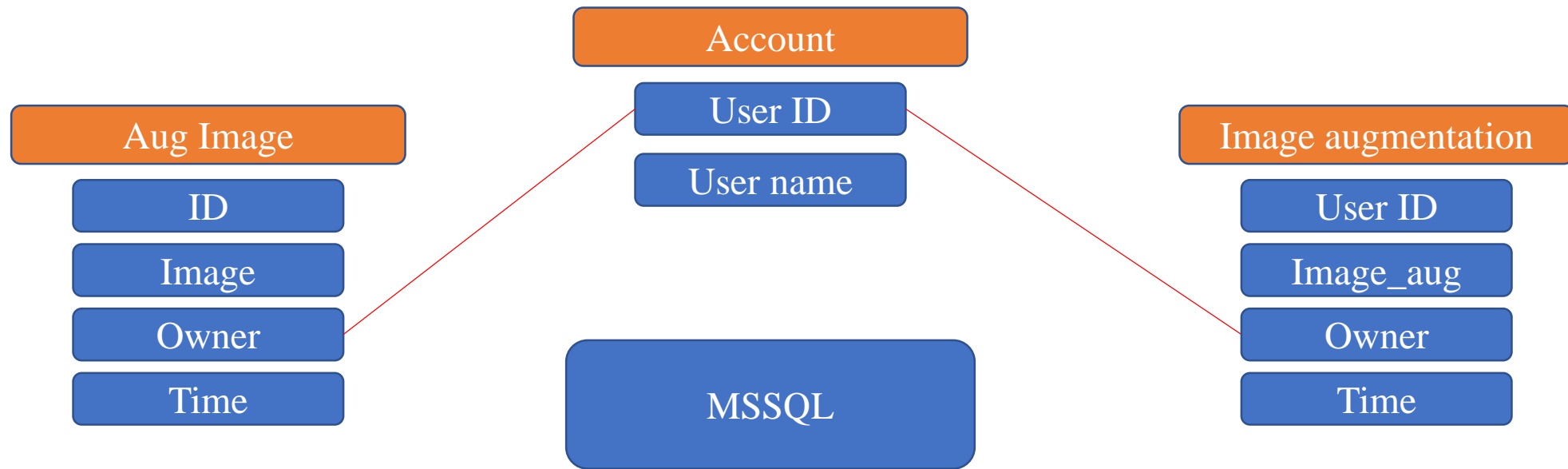
1

2

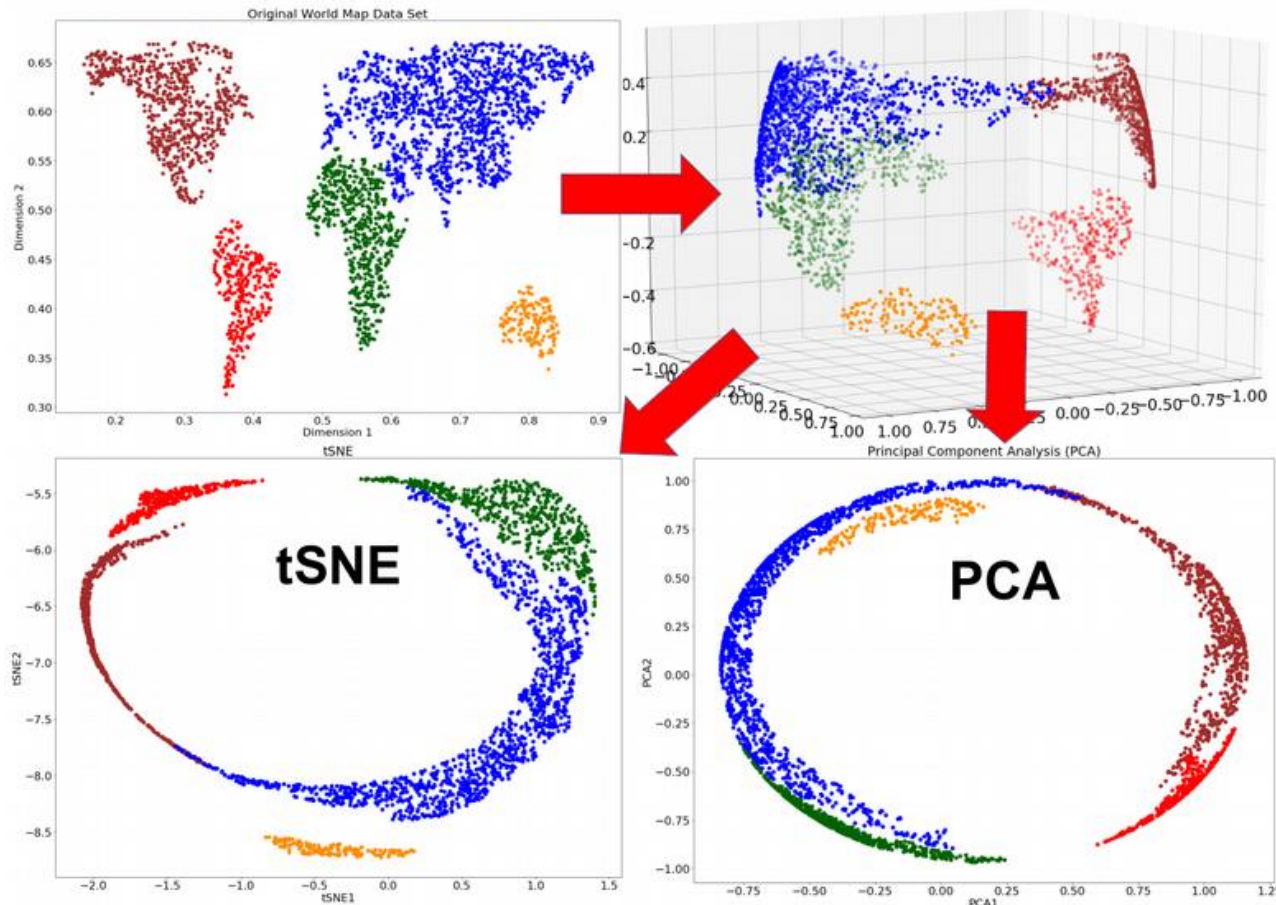


3

Image augmentation MSSQL

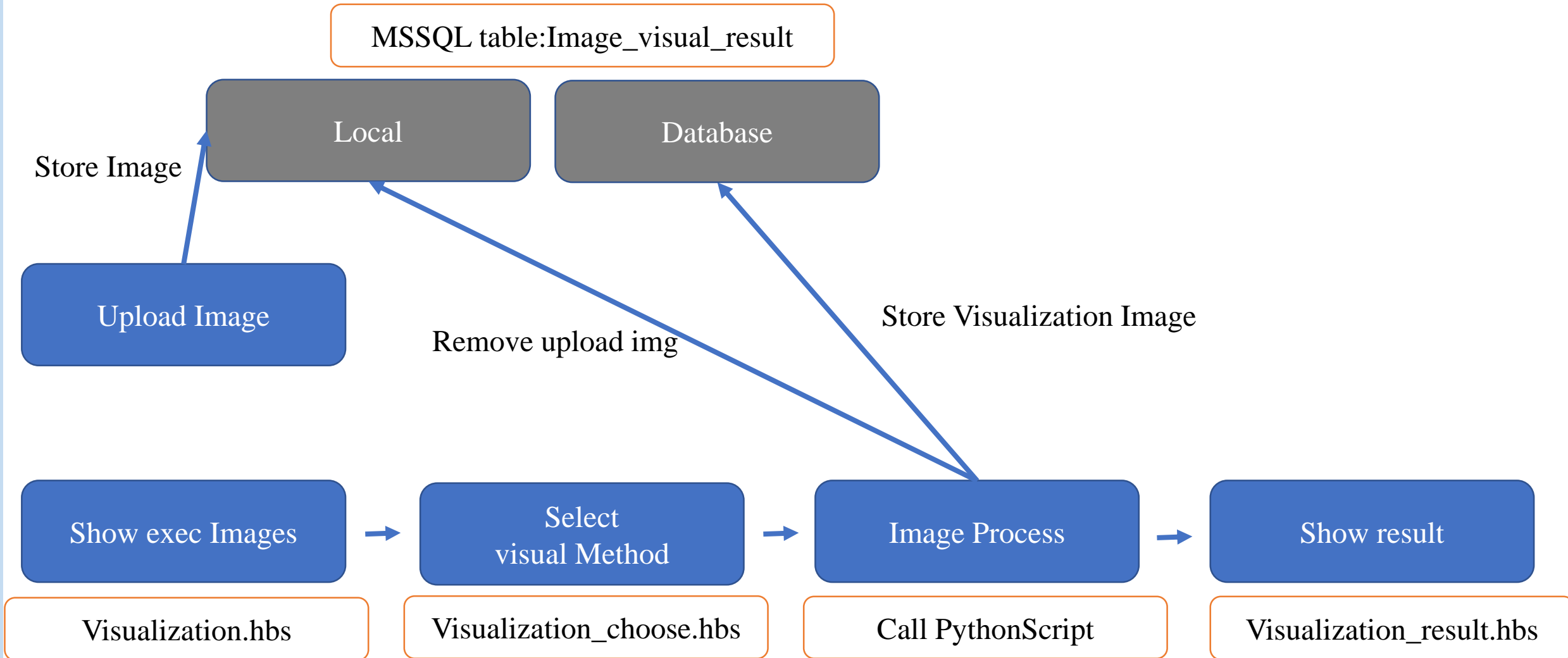


Visualization reduction of dimensionality



- Data cross validation(CV):
 - 同一分類的資料特徵會有所不同
 - 紋理
 - 背景
 - 角度...
 - 資料量較少的dataset進行隨機CV
 - 靠運氣
- 預期:
 - K-means分群設定
 - 視覺化降維協助資料挑選

Visualization Flow



Visualization



Upload Image

Show exec Images

Visualization.hbs

Auto
Visualization

資料降維視覚化
1.上傳資料集

降維視覚化
上傳資料集

upload

Upload multiple
images

選擇檔案 84檔案

Upload

AIIA LAB 2022

Visualization

降維設定

methods

Auto Visualization

資料降維視覺化
2.選擇PCA、t-SNE降維法
3.調整分群數量
4.執行

Select
visual Method

Visualization_choose.hbs

降維參數設定

降維方式: ☒ PCA ☐ t-SNE

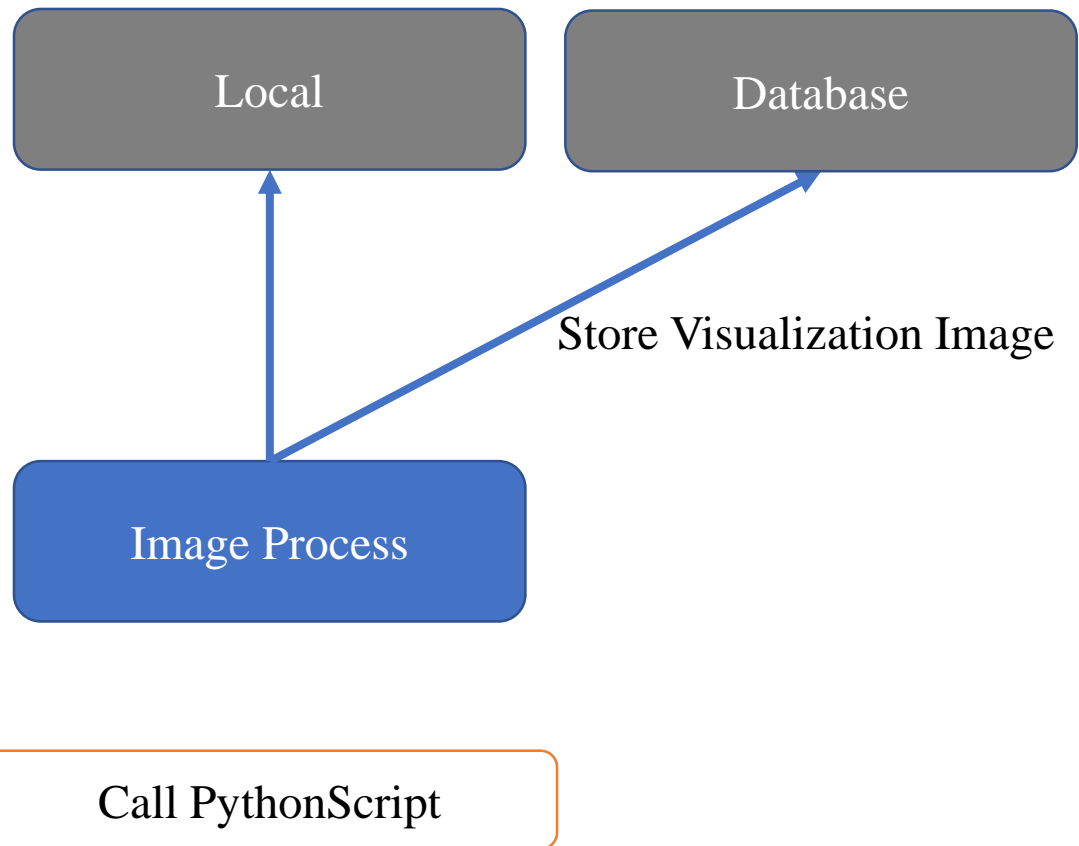
繪圖維度: ☒ 二維 ☐ 三維

分群數量:

submit

AIIA LAB 2022

Visualization



```
Visualization.py > ...
4  import sys
5  import random
6  import os
7  import shutil
8  import dbconfig
9  import datetime
10 from sklearn.cluster import KMeans
11 from sklearn.manifold import TSNE
12 from sklearn.decomposition import PCA
13 import matplotlib.pyplot as plt
14 from ImgProcessLib import foldImgResize
15
16
17
18 > def date_parser(): ...
28     #return str(microsecond)
29
30 > def prepareImgs(imagePath,reshape): ...
49
50 > def k_means(imgArr,cluster,method,savePath): ...
95
96 > def t_SNE(imagePath,components,reshape=False): ...
114
115 > def PCA_(imagePath,components,reshape=False): ...
```

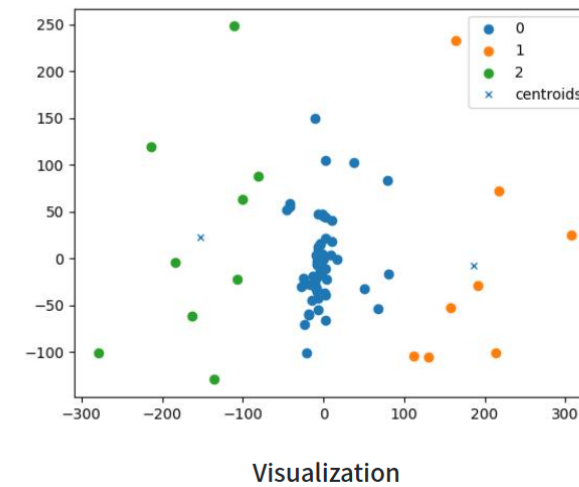

Visualization

Auto Visualization

資料降維視覚化
5. 找到最適合的分群協助資料挑選

Show result

Visualization_result.hbs



Visualization MSSQL

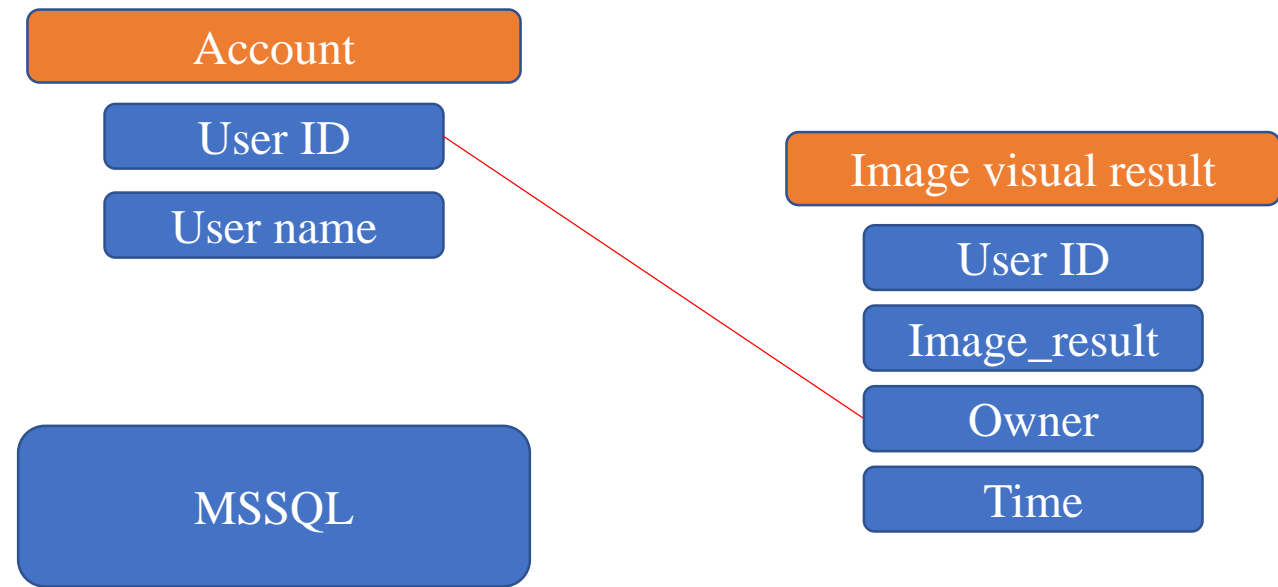
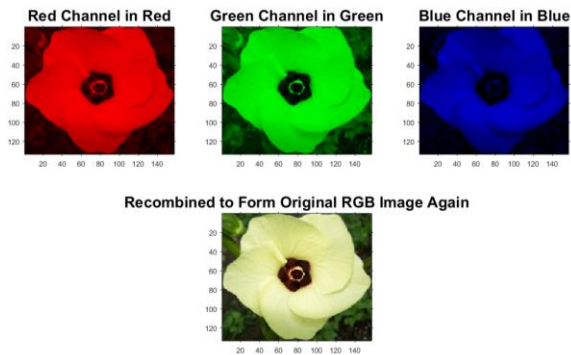
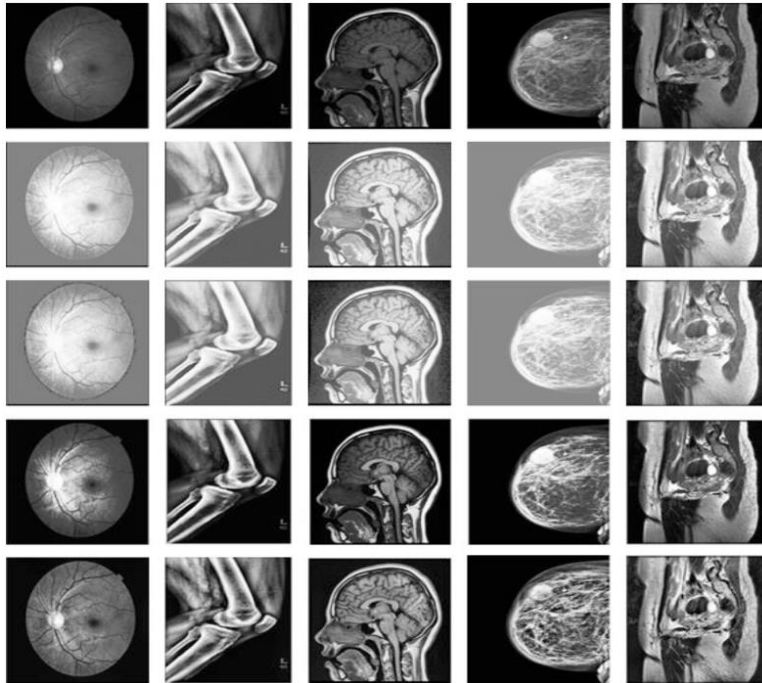


Image process



- 影像處理:
 - 使用傳統影像處理找出某些特徵
 - Histogram equalization
 - Split channels
 - Illumination equalization
 - ...

- 預期:
 - 快速觀察dataset某些特徵
 - 游標點選取得(R,G,B)

Label image



bounding box

- 需求功能:
 - Coedit
 - Voice
 - Schedule

- 資料標記:
 - bounding box & segmentation
- 目前軟體:
 - local & online (all existing)



segmentation