

NVLab Summer School 2023 Final Project

Group 2 members : 王浩、張睿鈞、蘇渝賢

Abstract

在本次Final Project 裡，我們參與AIdea 「AOI 瑕疵分類」(Defect Classifications of AOI)練習，根據任務目標、資料特性，自行設計模型classify 自動光學檢查的圖像，提升判讀瑕疵的分類效果。此外，我們也根據實驗，進而討論各種模型與資料的設計優劣。

1 Introduction

在「AOI 瑕疵分類」(Defect Classifications of AOI)的任務，處理的任務範圍是texture classification。在texture 分類任務常見的問題中，如果只使用CNN 的模型來進行分類，有可能會遇上無法同時capture 到local 與global 的information，因此我們在CNN 為基礎的模型上，接上Pyramid Pooling module，以獲取更全面完整的資訊。

有鑑於不少模型，在面對複雜的任務時，會使用ResNet、VGG 系列的模型作為Backbone，來extract features，因此本次Final Project 我們也有使用ResNet18 作為Backbone 取代相較簡單的Convolution layers 進行實驗，但可能因任務資料集的分佈與ResNet18 模型的Pre-train data 相差過大，因此結果分數較簡單的Convolution layers 低，最後不採納ResNet18 作為feature extractor。

除了Deep learning model，我們也使用處理圖像中常用到的統計方式來extract 出traditional feature，其中包含GLCM(GrayLevel Co-occurrence Matrix)與LBP(Local Binary Pattern)。最後再將Deep layer model extract 出來的feature 與Traditional feature 做fusion，以得出最後的分類結果。

2 Method

2.1 Framework

在圖1 可以看到，我們首先將input image 經過CNN 的model 作為Feature extractor，在經過兩個Module，分別為Traditional Feature Module 與Pyramid Pooling Module，各自代表擷取出圖像Texture的傳統演算法，與Deep learning extract 出來的features。並且最後將兩者做Fusion，接上Linear layer 得出最終的結果。

2.2 Pyramid Pooling Module

相較於常見的CNN-based model 如ResNet、VGG 只有一種scale size，我們參考PSPNet[ZSQ⁺17] 中的Pyramid Pooling Module，利用不同scale size 的pooling layer，來extract 出不同field 的information，在概念上可以包含到Local receptive field 與Global receptive field，增加模型的判別能力。

2.3 Traditional Feature Module

2.3.1 Gray-Level Co-occurrence Matrix

Gray-Level Co-occurrence Matrix(GLCM)是一種在影像處理和紋理分析中廣泛使用的統計方法，用於捕捉影像中像素強度的空間關係。GLCM是計算圖像中像素在空間中相對位置的灰階值出現的頻率，再透過一些屬性提出GLCM的紋理特徵。以下簡單介紹GLCM：

- 選取位置變化向量 $d(x,y)$ ： x 代表像素對之間的距離， y 代表像素對之間的角度。如圖2， $d(1,0)$ 代表夾角0度的長度1的位置變化向量，可統計8個角度，得到全方向GLCM。
- 計算GLCM矩陣：如圖3計算起始灰階值*i*到觀測灰階值*j*在原始圖片中出現的頻率，並將此值放置在GLCM矩陣第*i*列第*j*行的位置上。

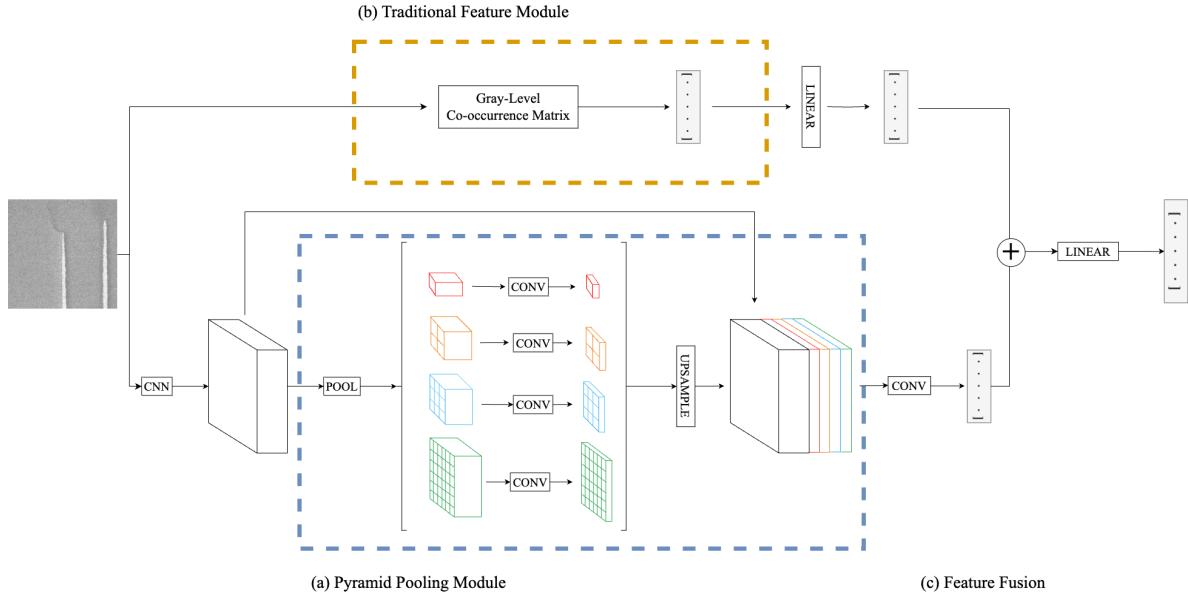


Figure 1: (a) Represent Pyramid Pooling Module, It can capture local and global features from different scale of pooling size (b) Traditional Feature Module can capture feature based on the GLCM and LBP algorithm.

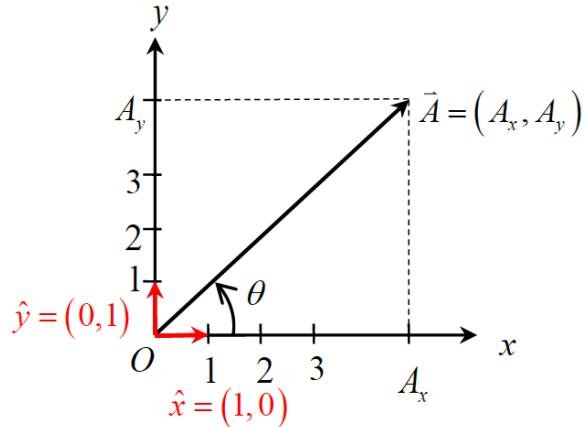


Figure 2: displacement vector

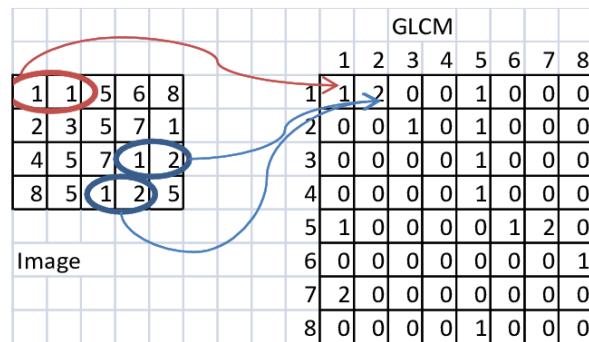


Figure 3: GLCM

- 提取紋理特徵：在GLCM矩陣上，透過一些描述紋理特徵的統計屬性，例如：Energy、Entropy、Contrast等，計算出紋理特徵。

2.4 Feature Fusion

根據文獻[Liu13]，可以得知，在面對Texture recognition時，我們可以利用已知的傳統方式，包含上述提到的Gray-Level Co-occurrence Matrix，並且串接Deep learning model 再進行Texture 相關任務。

另外我們得知文獻[LD18]，在面對較為複雜的Medical image 判讀時，一樣會先利用GLCM 來Extract 出texture 的資訊，並且再與Deep learning extract 出來的feature 做fusion，來得出精確的判斷。因此我們本次任務也採用類似架構進行實作。

3 Experiment

我們可以根據實驗後的表1 得知，分數最高為**0.9815**。而以下分別為不同的feature 討論。

Table 1: All result

| Backbone | | PPM | Augmentation | | | GLCM | LBP | performance |
|----------|--------|-----|--------------|------|-----------------|------|-----|---------------|
| CNN | ResNet | | flip | blur | color transform | | | |
| ✓ | | | | | | | | 0.9418 |
| ✓ | | | ✓ | | | | | 0.9701 |
| ✓ | | ✓ | | | | | | 0.9667 |
| ✓ | | ✓ | ✓ | | | | | 0.9802 |
| ✓ | | ✓ | ✓ | ✓ | | | | 0.9775 |
| ✓ | | ✓ | ✓ | | | ✓ | | 0.9815 |
| ✓ | | ✓ | ✓ | | | ✓ | ✓ | 0.9780 |
| ✓ | | ✓ | ✓ | | ✓ | ✓ | ✓ | 0.9664 |
| | ✓ | ✓ | ✓ | | | | | 0.9420 |

3.1 Model with/without PPM

我們首先實驗Pyramid Pooling Module 添加後的分數變化，可以從表2上看出，在加上PPM 之後，模型分數上升了0.01。

Table 2: Compare score about PPM

| Backbone | | PPM | Augmentation | | | GLCM | LBP | performance |
|----------|--------|-----|--------------|------|-----------------|------|-----|---------------|
| CNN | ResNet | | flip | blur | color transform | | | |
| ✓ | | | ✓ | | | | | 0.9701 |
| ✓ | | ✓ | ✓ | | | | | 0.9802 |

3.2 ResNet18 backbone models

有鑑於我們觀察其他研究[ZSQ⁺17, TZS⁺22]，發現深度學習模型往往在面對複雜任務時，會使用到ResNet系列或VGG 系列模型作為Backbone 來當成Encoder 或是Feature extractor，而我們此次根據參數量，也使用ResNet18 作為Backbone 進行實驗。但可能因為Defect Classifications of AOI 的任務資料集，與ResNet18 pretrain dataset 的分佈與性質有所差距，因此本次訓練在相同設定下，ResNet18 作為Backbone 的效果並沒有提升模型能力，分數可見表3。

Table 3: Compare scores with ResNet

| Backbone | | PPM | Augmentation | | | GLCM | LBP | performance |
|----------|--------|-----|--------------|------|-----------------|------|-----|---------------|
| CNN | ResNet | | flip | blur | color transform | | | |
| ✓ | | ✓ | ✓ | | | | | 0.9802 |
| | ✓ | ✓ | ✓ | | | | | 0.9420 |

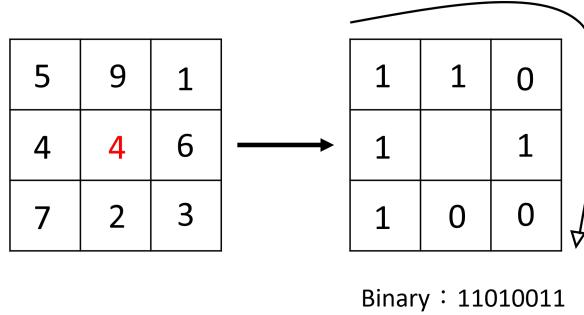


Figure 4: Illustration of LBP

3.3 Local Binary Pattern in Traditional Feature Module

LBP是一種用於圖像處理和紋理分析的feature extraction method，用來描述圖像中的local texture feature。對於每個pixel，LBP會在某個window範圍內將其周圍的pixel value和中心的pixel value來進行比較。如果周圍的值大於或等於中心的值，則該位置被標記為1，否則為0，最後將這些binary串在一起就形成了LBP code(如圖4)。

可見圖5，且根據[Liu13]提到，GLCM 與LBP 模型彼此在大部分環境條件下沒有「互補性」，可以參考圖6、沒有辦法降低特徵抽取後的缺點。因此我們也能夠從本次的實驗數據表4得知，在Traditional Feature Module 新增LBP 後，發現分數下降。

3.4 Data augmentation

- 我們觀察圖7 可以發現，資料集的defect 分佈尤其是Type 2, 3, 4 都是以垂直、水平軸為對稱，代表我們可以使用Horizontal、Vertical Flip 來作為Data Augmentation 的方式。根據表5，可以發現分數有顯著的提升。
- 此外，如果我們實驗採用Gaussian Blur 來擴增資料集，改動Image 的解析度，發現會降低模型的分數，可能因為在不同的image type 中，解析度也是資料的重要特徵之一，因此更動會產生負面的影響，可見表6。

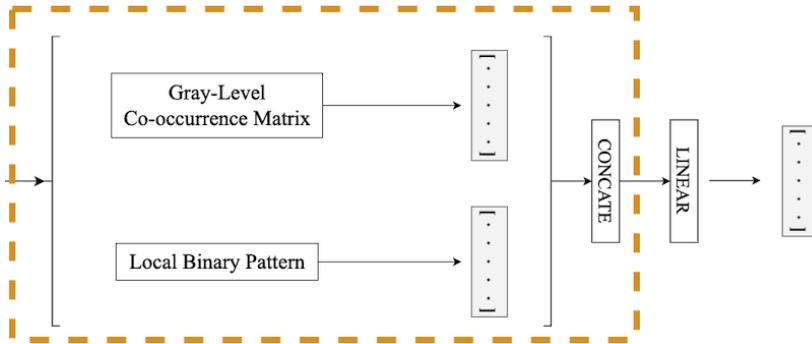


Figure 5: Traditional Feature Module can capture feature based on the GLCM and LBP algorithm.

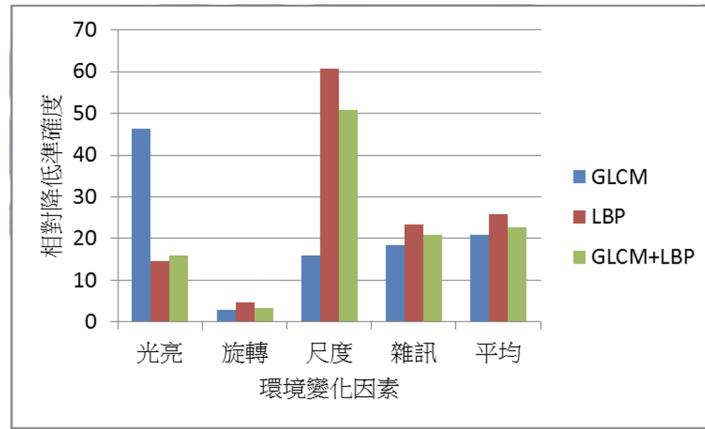


Figure 6: GLCM and LBP relatively reduce precision faced with different circumstance [Liu13]

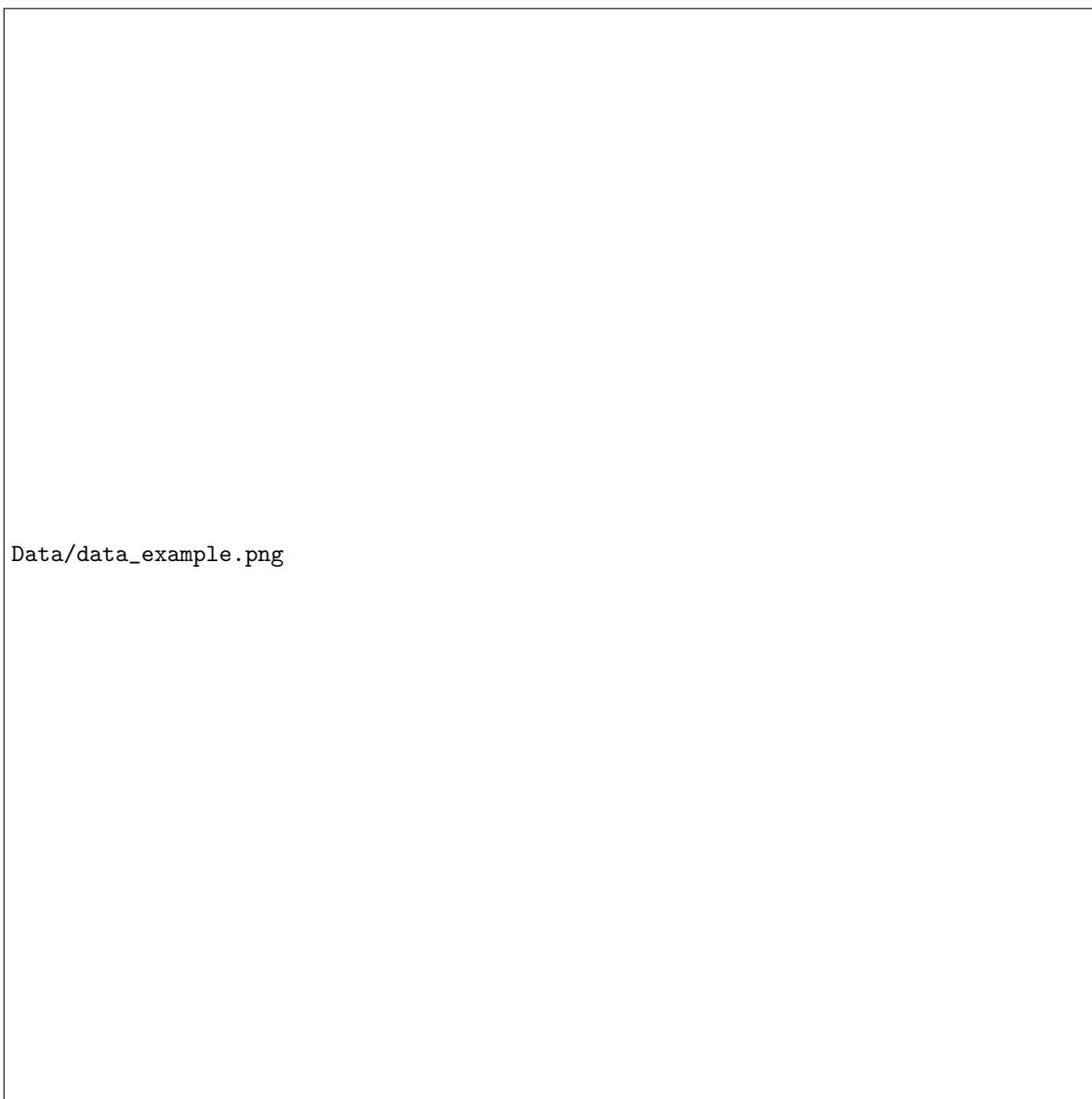


Figure 7: Every type of data example in training dataset

Table 4: Compare score about traditional features

| Backbone | | PPM | Augmentation | | | GLCM | LBP | performance |
|----------|--------|-----|--------------|------|-----------------|------|-----|---------------|
| CNN | ResNet | | flip | blur | color transform | | | |
| ✓ | | ✓ | ✓ | | | | | 0.9802 |
| ✓ | | ✓ | ✓ | | | ✓ | | 0.9815 |
| ✓ | | ✓ | ✓ | | | ✓ | ✓ | 0.9780 |

- 最後，我們也使用顏色的色階對比度的變化來作為實驗，但結果見表7，發現分數較低，可能是因為Model 對於灰度的高低判斷，也會因為顏色的轉換後失真，導致結果較差。

Table 5: Compare score about flip

| Backbone | | PPM | Augmentation | | | GLCM | LBP | performance |
|----------|--------|-----|--------------|------|-----------------|------|-----|---------------|
| CNN | ResNet | | flip | blur | color transform | | | |
| ✓ | | ✓ | | | | | | 0.9667 |
| ✓ | | ✓ | ✓ | | | | | 0.9802 |

Table 6: Compare score about blur

| Backbone | | PPM | Augmentation | | | GLCM | LBP | performance |
|----------|--------|-----|--------------|------|-----------------|------|-----|---------------|
| CNN | ResNet | | flip | blur | color transform | | | |
| ✓ | | ✓ | ✓ | | | | | 0.9802 |
| ✓ | | ✓ | ✓ | ✓ | | | | 0.9775 |

4 Future Direction

因為在這次project任務中的一張圖片只有一種類型，但如果之後的任務上，一張圖片可能包含了許多的種類，那我們的模型可能就必須再去進一步的修改，學習如何正確的分割出不同種類，而這也會讓模型在運用的範圍上更廣也更有效率。

如果texture classification model可以做出準確的判斷，或許可以解決其他領域的問題，像是透過衛星來幫助監控每年綠被的土地佔有比率來避免土地的過度開發，又或是可以監測在國家內是否過度的種植某種植物，導致物價的不平衡，在醫學上，我們也可以透過這種方式來分辨出正常細胞和病變細胞的差別，如果可以透過這種方法來標記出病變細胞，或許可以透過藥物的選擇性作用來減少對於正常細胞的損害。

5 Conclusion

我們在本次計畫中提出了一種texture分類任務的模型結構，透過使用PPM和GLCM的方法來提取更多的特徵，還有data augmentation的方法，比起單純使用傳統CNN我們的方法會更為準確。雖然我們只用在了一對一的任務上，但我們相信這種方法也可以應用於一對多的任務上。

References

- [LD18] ZhiFei Lai and HuiFang Deng. Medical image classification based on deep features extracted by deep model and statistic feature fusion with multilayer perceptron. *Computational Intelligence and Neuroscience*, 2018:2061516, Sep 2018.

Table 7: Compare score about color transformation

| Backbone | | PPM | Augmentation | | | GLCM | LBP | performance |
|----------|--------|-----|--------------|------|-----------------|------|-----|---------------|
| CNN | ResNet | | flip | blur | color transform | | | |
| ✓ | | ✓ | ✓ | | | ✓ | ✓ | 0.9780 |
| ✓ | | ✓ | ✓ | | ✓ | ✓ | ✓ | 0.9664 |

- [Liu13] Yen-Chen Liu. Improving the accuracy of texture recognition by multiple features. [碩士論文。國立交通大學] 臺灣博碩士論文知識加值系統, 2013.
- [TZS⁺22] Zhuotao Tian, Hengshuang Zhao, Michelle Shu, Zhicheng Yang, Ruiyu Li, and Jiaya Jia. Prior guided feature enrichment network for few-shot segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(2):1050–1065, 2022.
- [ZSQ⁺17] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia. Pyramid scene parsing network. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6230–6239, 2017.