

# Video Shot Change Detection

## Multimedia Content analysis

Wen-Hai Tseng

RE6124027

Nation Cheng Kung University

RE6124027@gs.ncku.edu.tw

### I. 程式執行環境

本次 Video Shot Change Detection 的透過使用 Siamese Network 作為 Visual Feature 並比較前後圖片的 Segment Map 差異。程式所需要使用的套件可以透過 requirements.txt 進行安裝，需要執行的程式檔為 predict.ipynb 檔，並且程式參考 <https://github.com/bubbliiiing/Siamese-pytorch> 進行修改。

### II. VISUAL FEATURE

影片拍攝變化檢測是影片處理領域中的一個重要問題，對於自動影片分割、事件檢測和內容檢索等應用具有關鍵意義。在本節中，我們將探討用於影片拍攝變化檢測的視覺特徵，這些特徵可用於檢測相鄰幀之間的變化，例如場景轉換、物體運動等。

#### A. Relate Work---顏色直方圖 (Color Histograms)

顏色直方圖是一種描述影像顏色分佈的統計特徵。通過計算每一幀的顏色直方圖，我們可以比較相鄰幀之間的色彩變化。例如，突然出現或消失的明顯色彩差異可能意味著場景變化。

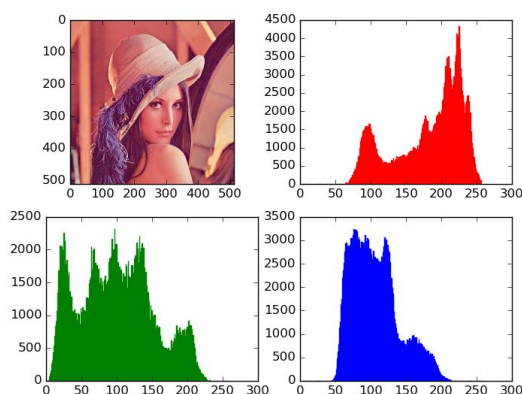


Fig.1 顏色直方圖範例.

#### B. Relate Work---邊緣檢測 (Edge Detection)

邊緣檢測是一種常用的影像處理技術，用於檢測影像中的邊緣結構。相鄰幀之間的邊緣變化可以指示著物體的運動或場景的變化。



Fig.2 邊緣檢測範例

Fig.4 ORB uses the BRIEF algorithm to compute descriptors for keypoints. The core idea of the BRIEF algorithm is to select a certain pattern of N point pairs around the keypoint P and combine the comparison results of these N point pairs as the descriptor.

#### C. 本次作業的做法--- Siamese neural network

孿生神經網路 (Siamese neural network)，又稱雙生神經網路，是一種基於兩個人工神經網路構建的耦合架構。孿生神經網路接受兩個樣本作為輸入，並輸出它們在高維空間中的嵌入表示，以比較這兩個樣本的相似程度。狹義的孿生神經網路由兩個結構相同且權重共享的神經網路拼接而成。廣義的孿生神經網路，或稱“偽孿生神經網路 (pseudo-siamese network)”，則可以由任意兩個神經網路拼接而成。孿生神經網路通常具有深度結構，可以由卷積神經網路、循環神經網路等組成。

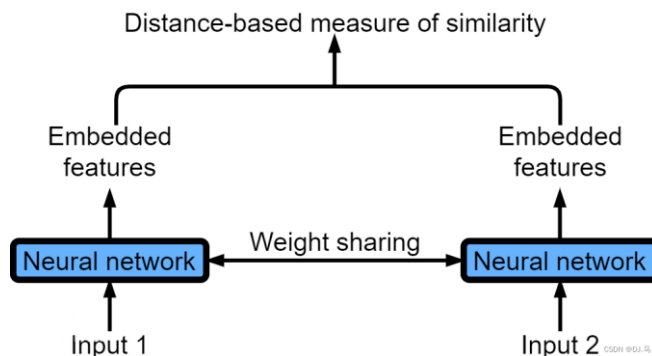


Fig.3 孿生神經網路流程

我選擇了使用圖像分割模型 `deeplabv3_resnet101` 來提取視覺特徵。這個模型具有深度學習架構，能夠對圖像進行高效的像素級別的分割，即將圖像中的每個像素分配到相應的類別中。通過這種方式，我可以獲得更具信息量的特徵來比較圖片的相似度。

使用 `deeplabv3_resnet101` 模型的好處在於它不僅能夠識別圖片中的對象和場景，還能夠捕捉細微的特徵和紋理。這對於需要精確的視覺特徵來比較相似度的任務非常重要。

通過這個方法，我可以將圖片轉換為具有高維度的特徵表示，這些特徵可以更有效地捕捉到圖片的結構和內容。藉由將兩張圖片的特徵表示進行比較，我可以計算它們之間的相似度，進而判斷它們之間的關聯程度。

總的來說，使用 `deeplabv3_resnet101` 模型進行視覺特徵提取的方法能夠幫助我更全面、準確地比較圖片之間的相似度，從而應用於各種需要圖像相似度比較的任務中。

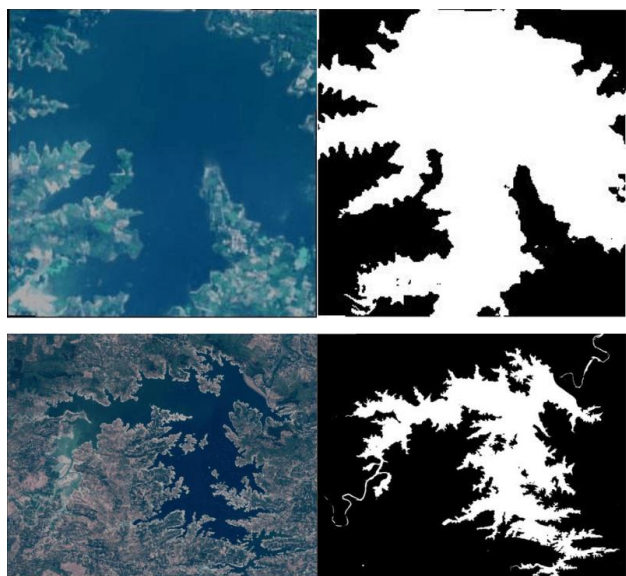


Fig.4 `deeplabv3_resnet101` 模型進行視覺特徵提取

### III. SHOT CHANGE DETECTION 演算法

我使用了 SHOT CHANGE DETECTION 演算法，該演算法基於深度學習模型 `deeplabv3_resnet101` 來提取視覺特徵。首先，我將每隔 20 張圖片提取的特徵圖進行平均差計算，以衡量這些圖片之間的變化量。如果平均差大於 0.7，則判定這 20 張圖片之間可能存在場景切換。

接下來，我需要進一步檢查這 20 張圖片中每張圖片的前後變化量。如果只有一張圖片的變化量大於 0.7，則將其判定為場景切換。如果有多於一張圖片的變化量大於 0.7，則將這些圖片都判定為場景切換。

如果 20 張圖片間沒有任何變化量大於 0.7 的圖片，且首尾兩張圖片的變化量都小於 0.9，這可能表示著一個變化度較慢的場景切換效果，因此也需要將其檢測出來。

總的來說，這個演算法利用了深度學習模型提取的視覺特徵來檢測影片中的場景切換，並根據特徵間的變化量和分佈模式進行判斷，以確定場景切換的位置和數量。

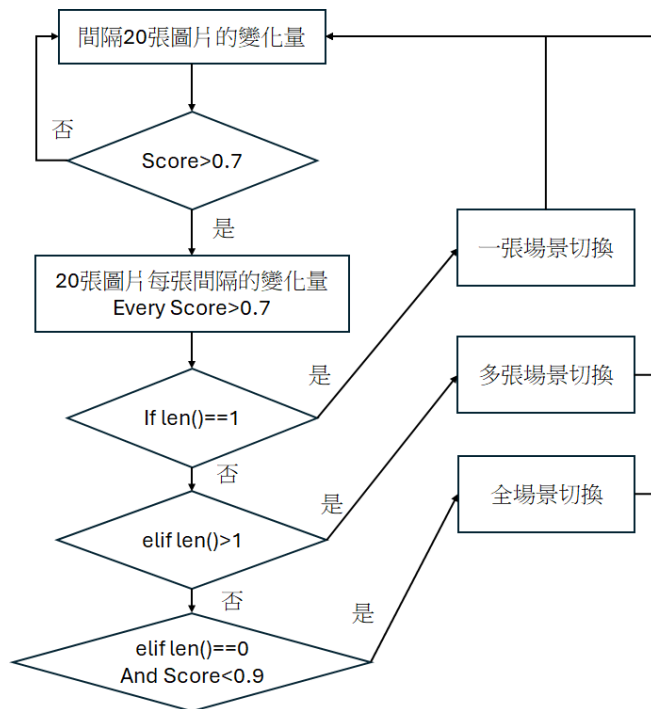


Fig.5 SHOT CHANGE DETECTION 流程

### IV. 偵測效能

在這次的實驗中，我採用了 PR 曲線（Precision-Recall Curve）、混淆矩陣和準確度作為檢測模型性能的標準。

首先，PR 曲線是一種用於評估二元分類器性能的常用工具。它以查全率（Recall）為橫軸，精確率（Precision）為縱軸，對模型的性能進行評估。PR 曲線展示了在不同閾值下的模型的查全率和精確率之間的平衡關係。在這個實驗中，我將根據 PR 曲線的形狀和曲線下面積（Area Under Curve, AUC）來評估模型的性能，以判斷模型對於場景切換的檢測效果。

其次，我還使用混淆矩陣作為評估指標。混淆矩陣是一個二維的表格，用於顯示模型的預測結果與實際標籤之間的對應關係。在這個案例中，混淆矩陣將顯示模型將每個類別（場景切換和非場景切換）的實際預測情況，有助於評估模型的準確性、召回率和特異性等性能指標。

最後，準確度（Accuracy）是另一個常見的模型性能指標，它表示模型對於所有樣本的預測準確率。在這

次實驗中，我將使用準確度來評估模型整體的性能表現。

	climate	news	NGC
ACC	0.9449	0.95	0.67

Table2. 三個資料集的準確度

綜合 PR 曲線和混淆矩陣的分析，我們可以觀察到與其他資料集相比，對於 news 資料集的場景切換檢測效果明顯較差。儘管準確度指標顯示整體性能相對較好，但從 PR 曲線和混淆矩陣來看，我們發現在這個資料集上模型的表現並不理想。

進一步分析發現，我們的方法在檢測慢速切換效果方面表現良好。這意味著我們的模型能夠成功檢測出變化度較慢的場景切換，這在某些應用場景下可能是非常有用的。

然而，我們也發現在單張圖片之間的順間切換檢測中存在著較高的誤檢率，並且產生了較多的 false alarm（誤報）。這可能是因為模型對於單張圖片之間微小變化的敏感性較高，導致了一些不必要的切換檢測。

綜上所述，雖然我們的方法在檢測慢速切換方面有較好的效果，但對於單張圖片之間的順間切換仍存在改進的空間。這可能需要更進一步的調整模型參數或引入更多的後處理技術來改善檢測精度和降低 false alarm 率。

REFERENCES

[1]<https://steam.oxxostudio.tw/webp/python/ai/opencv-edge-detection-03.webp>  
[2][https://lijiancheng0614.github.io/2016/03/14/2016\\_03\\_14\\_Color\\_Histogram/figure\\_RGB.png](https://lijiancheng0614.github.io/2016/03/14/2016_03_14_Color_Histogram/figure_RGB.png)  
[3] <https://blog.csdn.net/DJJ5210/article/details/133203440>  
[4]<https://debuggercafe.com/wp-content/uploads/2023/01/train-pytorch-deeplabv3-on-custom-data-gt-image-pairs.png>

IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being p

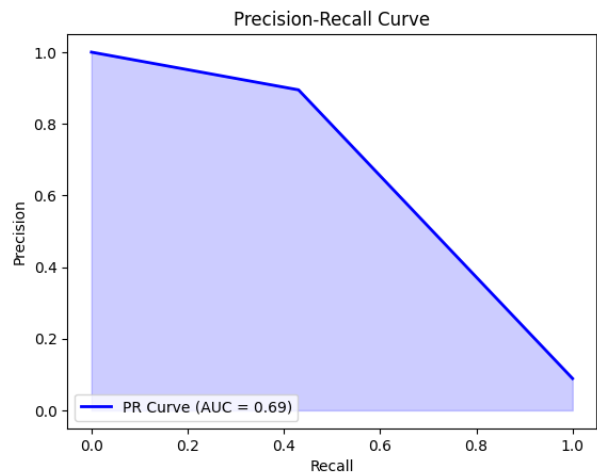


Fig.5 climate PR Curve

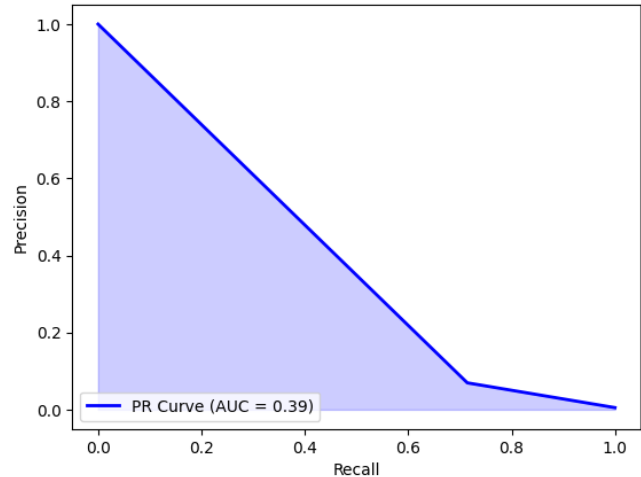


Fig.6 news PR Curve

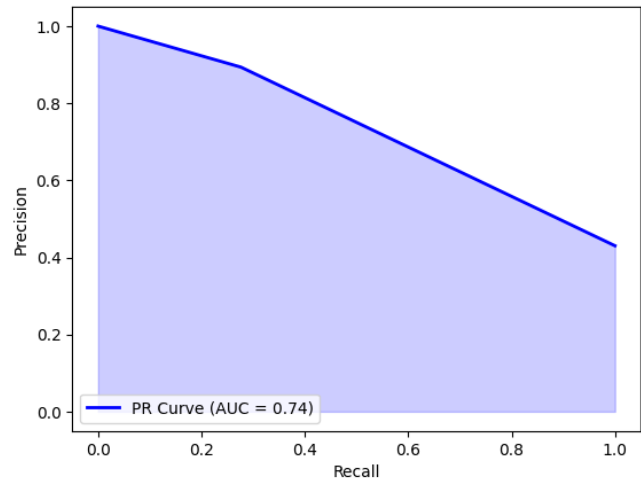


Fig.7 NGC PR Curve

climate		news		NGC	
1614	8	1306	67	589	15
90	68	2	5	330	126

Tabell. 三個資料集的混淆矩陣