

## 專題成果報告

主題：應用深度學習技術作黑白漫畫上色之研究

組員：許志仲、黃于九

指導教授：余執彰

## 一、研究動機

至今已經有許多AI圖像上色工具可供使用，但依舊無法完全仰賴於機器上色，而這也是本專題所聚焦改善的方向，目標能夠對傳統的日式黑白漫畫進行全自動上色（Automatic Colorization）。其中本研究要處理的問題大致如下：

1. 訓練用的資料格式不一，如何對資料進行預處理。
2. 探討要如何運用有限的數據量來進行深度學習。
3. 自動上色方式的神經網路（Neural Net-work）之架構設計。
4. 如何對日式漫畫的特徵做個別處理，包含網點、對話框、擬聲詞、漸層，這些技法的上色方式較為特殊，需要額外做處理。

## 二、研究方法與結果

以下為本專題的上色流程設計：首先輸入原始漫畫，將其轉換成線稿後，接著輸入給線稿轉灰階的上色模型，將得到的灰階漫畫再輸入給灰階轉彩色的上色模型，最後會得到彩色漫畫。為了簡潔說明，以下會將線稿轉灰階的上色模型、灰階轉彩色的上色模型分別稱作 Model1、Model2。

### 第一部分：資料集處理

本專題使用「海賊王漫畫彩色版本」作為彩色漫畫資料集（共計 12477 張圖片），進一步對原圖片套用灰階轉換，取得灰階漫畫資料集，再藉由高斯模糊找到圖片的輪廓，其運算結果作為線稿漫畫資料集。因為是採用監督式學習，我們在製作資料集時，針對不同資料夾下的圖片使用相同的命名，確保圖片之間正確對應。

#### Step 1

將原圖片縮放成 256\*256 像素，提高模型的收斂速度。

#### Step 2

將圖像從轉換成 LAB 模式。對於 Model1 來說，圖片以 L 通道數值作為輸入，會得到輸出 L' 數值並成為 Model2 的輸入，最後會輸出預測的 AB 通道數值。因此我們可以藉由輸入圖片提供的資訊，預測上色。

#### Step 3

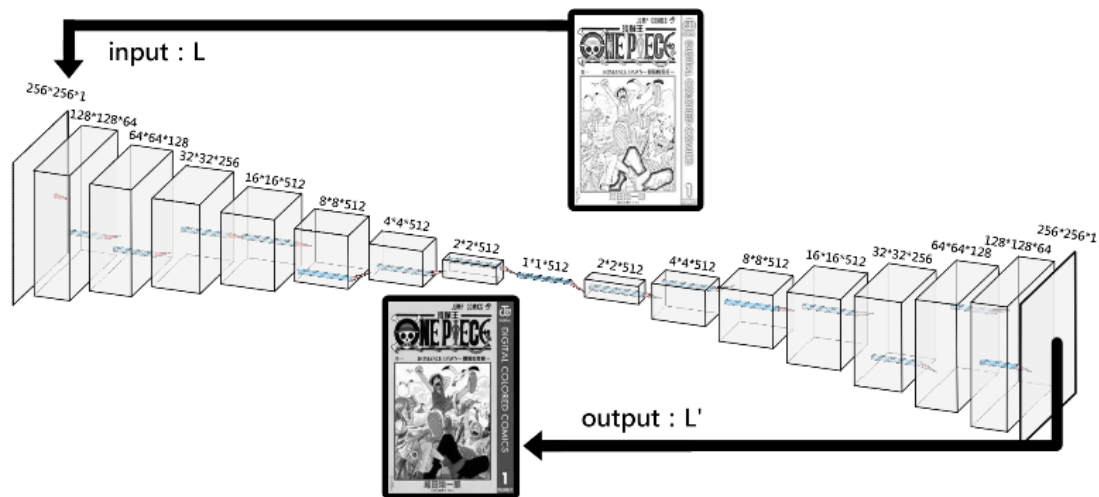
為了讓模型更有效學習漫畫物件特徵，在模型訓練過程中會對圖片進行 shuffle，可以防止模型抖動，避免發生過擬合，並且使得模型學到更加正確的特徵。

### 第二部分：模型架構說明

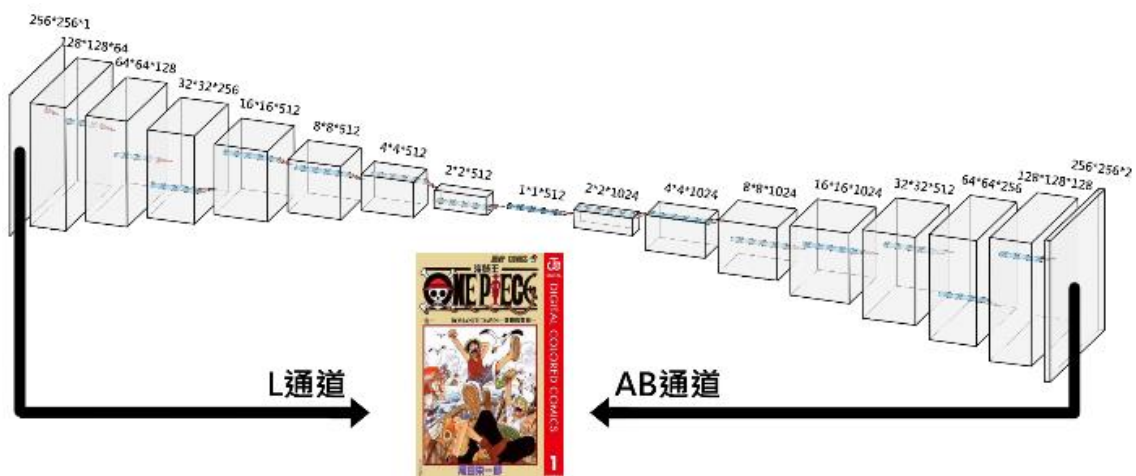
本專題中的 Model1 與 Model2 都是採用 GAN（Generative Adversarial Network）來建立模型架構，包含生成器（Generator）、判別器（Discriminator）。

#### 生成器：

Model1 和 Model2 的生成器如圖（一）、圖（二）所示，參考 U-net 架構設計而成。對 Model1 生成器的輸入一張 256\*256 的單通道（L）線稿圖片，經由卷積層（Convolution Layer）後，會獲得 64 張 128\*128 的特徵圖（Activation Maps），依此類推。到達 U-net 最底層時，共有 512 張 1\*1 的特徵圖，由於 U-net 是對稱結構，因此還會再進行多次反卷積運算，輸出一張 256\*256 的單通道（L'）灰階圖片；Model2 生成器也是依照相同的模式搭建出來的，不同之處在於 Model2 的輸入是 Model1 的輸出，而 Model2 的輸出則是雙通道（AB）彩色圖片，最後再將 L' 通道與 AB 通道合併，會得到上色成果圖。



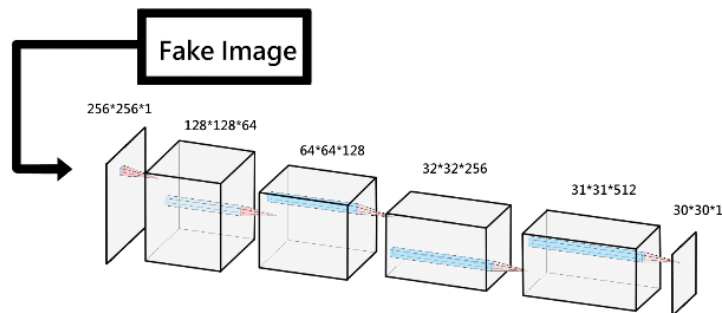
圖（一）Model11 生成器架構圖



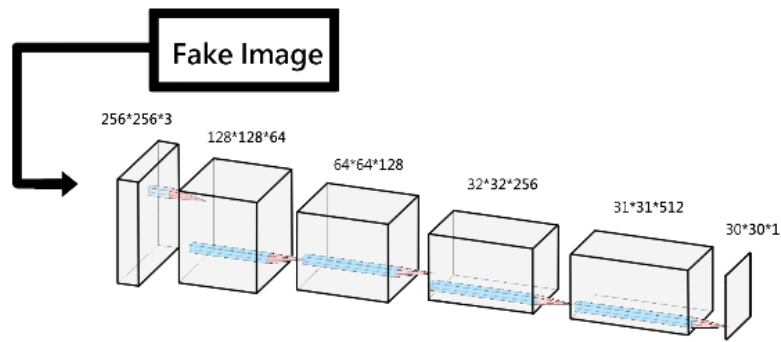
圖（二）Model12 生成器架構圖

### 判別器：

Model11 和 Model12 的判別器如圖（三）、圖（四）所示，我們使用 PatchGAN 來鑑定圖像的生成品質，其輸出為一張  $N \times N$  的特徵圖，基於感受野（Receptive Field）在圖像上的預測結果來計算損失（Loss），能夠考慮到圖像不同部分的影響，讓模型能夠更關注圖像細節，對上色也較有幫助。Model11 判別器的輸入為 Model11 生成器所輸出的灰階圖片，經過一連串的卷積運算後，會輸出一張  $30 \times 30$  的特徵圖，圖中的每一個 Patch 會用一個介於 0 到 1 之間的數值表示，1 為真，0 為假，再與真實灰階圖片對照，進行評估。Model12 判別器的設計概念與 Model11 判別器相同，唯一不同之處在於 Model12 判別器的輸入為 Model12 生成器所輸出的彩色圖片，其輸出會與真實彩色圖片對照，評估結果好壞。

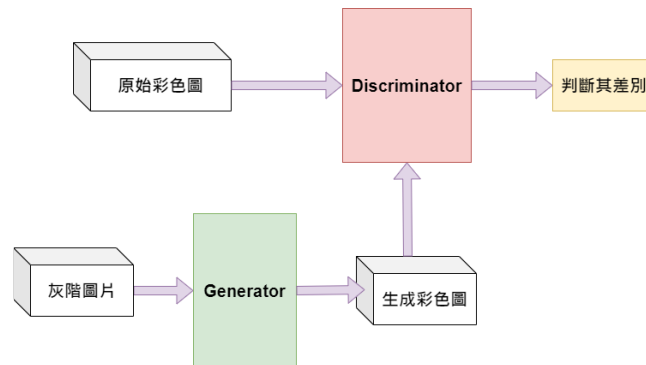


圖（三）Model11 判別器架構圖



圖（四）Model2 判別器架構圖

接著要來探討生成器與判別器之間是如何互動的，以下將以 Model2 的生成器和判別器來舉例說明。生成器會訓練出一個神經網路，用來產生彩色圖，而判別器會判定生成彩色圖與原始彩色圖的差別，分辨圖片是生成彩色圖或是原始彩色圖，並且為該圖片評分，示意如圖（五）。要成功騙過判別器的話，生成器必須產出以假亂真的彩色圖片。

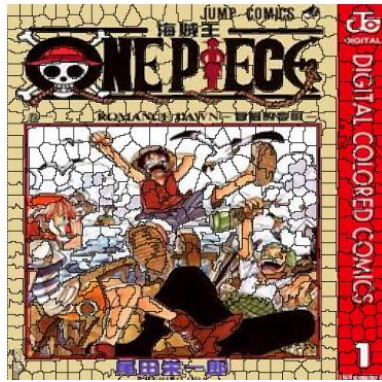


圖（五）Model2 訓練流程圖

### 第三部分：分析與優化

為了使模型上色更豐富多樣，我們基於現有的模型架構，嘗試將Model1和Model2的生成器架構改以Resnet34建構成的U-net取代，並分別訓練1500個Epoch及700個Epoch，而後引入結構相似性指標（structural similarity index measure, SSIM），將原始彩色圖片與生成彩色圖片作為輸入，輸出SSIM數值（介於0到1之間），從而判斷其相似程度。我們從資料集中隨機取樣100張灰階圖片進行上色後，再與對應的原圖比較，平均得到的SSIM數值高達0.82，較原架構高出0.7，足以證明新的生成器架構具備更好的學習力。

在實驗過程中，察覺到某些生成彩色圖片的大面積色塊較不平均，視覺觀感上較為突兀，因此導入了SLIC（Simple Linear Iterative Cluster）演算法，用來生成超像素（Super-pixel）圖片來做處理，得到的像素分割如圖（六）所示，接著會計算同一類別的所有像素的色彩平均值，最後用色彩平均值取代該類別的所有像素點，如圖（七）所示。我們將生成彩色圖片轉換成超像素圖片並讓其取代，使區域色塊較為平均。



圖（六）超像素分割示意圖，K=1000



圖（七）色塊平均示意圖，K=1000

### 三、 結論與展望

我們成功透過二階段上色模型解決特定漫畫區域不被上色的問題，搭配 Resnet34 的 U-net 生成器架構，增添了更多上色可能性，之後結合 SLIC 演算法生成的超像素圖片，進行後處理，使得整體畫面較為協調。

相較於其他自動上色工具，本專題現階段已經能夠提供航海王漫畫一個良好的上色效果，上色模型準確學習到漫畫物件的特徵，像是漫畫對話框、頁邊與間隔通常不上色、天空與海洋會是藍色、人臉是皮膚色等。但由於不同漫畫的物件線條以及上色風格差異頗大，上色模型需要面對前所未見的資訊，導致上色成果不甚理想，因此，若要將模型運用在其他漫畫的上色工作上，仍然是一大挑戰。



圖（八）上色結果及展示介面

### 四、 未來展望

由於輸出圖片經過縮放，導致畫面解析度低及處理後的對話框文字模糊，下一步會嘗試加入文字區塊辨識以保留文字的原始樣式，至於圖片解析度低的問題可以運用超解析度還原的模型，使得圖片放大後可以保持較高的解析度。

## 五、 參考文獻

- [1] Paulina Hensman, Kiyoharu Aizawa. cGAN-based Manga Colorization Using a Single Training Image. arXiv:1706.06918v1 [cs.GR] 21 Jun 2017.
- [2] Gustav Larsson<sup>1</sup>, Michael Maire<sup>2</sup>, and Gregory Shakhnarovich<sup>2</sup>. Learning Representations for Automatic Colorization. arXiv:1603.06668v3 [cs.CV] 13 Aug 2017.
- [3] CHIE FURUSAWA\*, KAZUYUKI HIROSHIBA\*, KEISUKE OGAKI, YURI ODAGIRI, YURI ODAGIRI, DWANGO Co., Ltd. Comicolorization: Semi-Automatic Manga Colorization. arXiv:1706.06759v4 [cs.CV] 28 Sep 2017.
- [4] Yiyang Yan. ZING! POW! - COMIC COLORIZATION WITH CGAN. Fall 2018.
- [5] Yifan Liu <sup>1</sup> , <sup>2</sup> , Zengchang Qin <sup>1</sup> , Zhenbo Luo <sup>2</sup> , and Hua Wang <sup>2</sup>. Auto-painter: Cartoon Image Generation from Sketch by Using Conditional Generative Adversarial Networks. arXiv:1705.01908v2 [cs.CV] 7 May 2017.
- [6] 梁培俊, 刘怡俊. 基于条件生成对抗网络的漫画手绘图上色方法《计算机应用研究》2019 年第 36 卷第 2 期.
- [7] Ian J. Goodfellow, Jean Pouget-Abadie\* , Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair , Aaron Courville, Yoshua Bengio. Generative Adversarial Nets. arXiv:1406.2661v1 [stat.ML] 10 Jun 2014.

110 學年度科技部大專生研究計畫補助

計畫編號：110-2813-C-033-021-E

學生 黃于九

計畫為專題研究的主題，經討論後由另一位同學代表提出申請