

國立臺灣大學電機資訊學院資訊工程學系

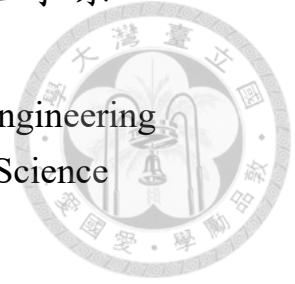
碩士論文

Department of Computer Science and Information Engineering

College of Electrical Engineering and Computer Science

National Taiwan University

Master Thesis



基於卷積神經網路與色彩感知的水彩混色模型

Perceptual-Based CNN Model for Watercolor Mixing
Prediction

黃雅博

Ya-Bo Huang

指導教授：歐陽明博士

Advisor: Ming Ouhyoung, Ph.D.

中華民國 107 年 6 月

June, 2018





國立臺灣大學碩士學位論文
口試委員會審定書

基於捲積神經網路與色彩感知的水彩混色模型

Perceptual-Based CNN Model for Watercolor Mixing
Prediction

本論文係黃雅博君（學號 R05922124）在國立臺灣大學資訊工程學系完成之碩士學位論文，於民國 107 年 6 月 22 日承下列考試委員審查通過及口試及格，特此證明

口試委員：

歐陽明

（指導教授）

梁容輝

傅楸善

莊永裕

系主任





誌謝

感謝我的家人，我的同學，在研究生活中一路相互支持。

特別感謝研究夥伴鎡璠學姊，以及歐陽教授，在研究上給我很多的幫助與啟發。

要感謝的人太多，最後感謝上帝。



Acknowledgements



This project was partially supported by Meidatek Inc. under Grant No.: MTKC-2018-0167, Ministry of Science and Technology(MOST), Taiwan under Grant No.: 106-3114-E-002-012 and 105-2221-E-002-128-MY2





摘要

本論文主要探討將卷積神經網路 (CNN) 應用於水彩混色預測的模型，以及利用色彩感知誤差作為損失函式對於模型訓練的影響。目前顏料混色之研究多以 Kubelka-Munk 理論作為基礎，然而 K-M 理論本身的限制，無法完全符合半透明的水彩顏料預測之需求。因此，設計一個更符合的水彩顏料的混色模型，能夠對未來水彩模擬以及相關應用有所貢獻。本文用一多層卷積神經網路來學習兩顏料混色後的反射頻譜。訓練資料部分使用 [Chen et al. 2018] 之水彩資料庫，訓練目標是最小化混色後的反射頻譜在上與真實混色之色彩誤差。透過本文所提出的模型以及損失函式，混色預測在測試資料中，有 88.7% 的結果能夠達到 $\Delta E_{Lab} < 5$ 的色彩誤差，也就是人眼無法輕易分辨其色差之程度。此外，本研究發現一發生在預測結果上之特別現象。此模型透過色感誤差的損失函式訓練後所產生的結果在頻譜上，相較於真實混色，會有中幅度的波動，無法貼近真實頻譜。然而在色彩上，卻因為損失函式，而能在色彩上與真實混色貼近。

關鍵字：顏色匹配函式、顏料混色、光譜、卷積神經網路、機器學習





Abstract

In the paper, we propose a model to predict the mixture of watercolor pigments using convolutional neural networks (CNN). With a watercolor dataset, we train our model to minimize the loss function of sRGB differences. In metric of color difference ΔE_{Lab} , our model achieves 88.7 % of data that $\Delta E_{Lab} < 5$ on the test set, which means the difference cannot easily be detected by the human eye. In addition, an interesting phenomenon is found; Even if the reflectance curve of the predicted color is not as smooth as the ground truth curve, the RGB color is still close to the ground truth.

Keywords: Color Matching Function , Pigment Mixing, Spectrum, Convolutional Neuron Network, Machine Learning





Contents

誌謝	v
Acknowledgements	vii
摘要	ix
Abstract	xi
1 Introduction	1
2 Related Work	3
2.1 Kubelka-Munk Model for pigment color mixing	3
2.2 Neuron Network for Color Mixing	5
2.3 CIE Color Matching Functions	7
2.4 Development of Convolutional Neuron Network	9
3 Method	11
3.1 Convolutional Network for Pigment Mixing	12
3.2 Perceptual Loss Function to Color	15
4 Experiments	21
4.1 Training Details	21
4.2 Result	25
5 Discussion	31

6 Conclusion

35

Bibliography

37





List of Figures

1.1	Part of color mixing results (a)~(f), showing different cases in our test set. With two input pigments PA and PB, our model can predict the results (Ours, left) which are almost indistinguishable to the ground truth (GT, right) of the real mixed pigment.	2
2.1	圖為 Wright 與 Guild 的實驗設計，觀測者利用可控亮度的三原色來匹配測試顏色。	6
2.2	CIE 1931 RGB 顏色匹配函數。顏色匹配函數是匹配水平刻度標示的波長的單色測試顏色所需要的原色數量。	8
2.3	圖例引用自 He et al. 之論文。顯示出無論在 training(左) 與 testing(右) 時，使用過深的”普通”神經網路會造成較高的誤差。	10
3.1	An illustration of the architecture of our model. We use an alternative ResNet18 to extract multi-level features from input reflectance and transmittance (RA, RB, TA, TB and Paper). A global average pooling layer is connected to the last conv. layer to reduce the number of parameters in our model. Fully connected layers after the pooling layer perform regression to minimize the perceptual loss between output and ground truth.	12
3.2	An illustration of our modified ResNet18 network, which is built with different filter sizes, pooling and output layers to generate a reflectance spectrum.	14
3.3	圖為在 380nm 至 780nm 之間的 CIE XYZ 標準觀察的色彩匹配函數。	16
3.4	Spectral power distribution of D65.	17

4.1	圖為 RGB loss 收斂曲線，可以看到損失收斂曲線到了 8k 次訓練後已經到達飽和狀態。	23
4.2	圖為 Delta E 收斂曲線，與同為基於 XYZ 空間的 RGB loss 有類似的收斂曲線。	24
4.3	圖為訓練後的模型在測試集合上的 ΔE_{Lab} 統計結果。	25
4.4	圖為訓練後的模型在測試集合上的 ΔE_{Lab} 的累積分布圖。	25
4.5	Delta E < 2 之部分結果	26
4.6	Delta E < 3.5 之部分結果	26
4.7	Delta E < 5 之部分結果	27
4.8	Delta E > 5 之部分結果	27
4.9	測試集中之 worst case	28
4.10	圖為訓練後的模型在測試集合上的 RGB 顏色上的結果。其中每個小色塊的左邊為 Ground Truth，右邊為預測出來的混色結果。結果由左至右，由上至下，依 ΔE_{Lab} 數值由小至大排列。	29
5.1	此圖說明：暗色調混色容易因反射光譜的雜訊與飄移而造成色差。左圖左邊為實際混色，右邊為預測結果。右圖藍色線為實際混色光譜，橘色為預測結果。	31
5.2	An illustration of the output reflectance of our model. The curve between visible wavelengths (380 to 750 nm) mainly focus on the short-term average, while the curve beyond 750nm can be totally ignored (being invisible to human eye)	32
5.3	此圖說明：白色混色在光譜上的偏重於三個區域。左圖左邊為實際混色，右邊為預測結果。右圖藍色線為實際混色光譜，橘色為預測結果。	33

List of Tables







Chapter 1

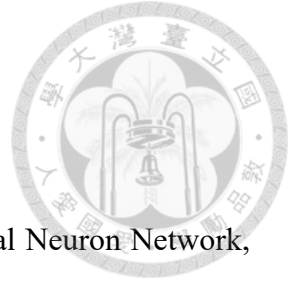
Introduction

顏料混色的模擬在數位繪畫中扮演著關鍵的角色之一，能使藝術家作數位繪畫時得到更加正確且合理的結果。而設計良好的混色模型能夠幫助數位繪畫中的色彩呈現更加真實與精確。

先前的顏料混色的研究者如 Haase and Meyer [11] 與 Baxter et al. [3] 都使用了 Kubelka-Munk [16] 的方法來建立混色模型，透過顏料原色之散射值與吸收值做線性組合，來計算顏料混色後的散射值與吸收值。然而，有研究指出 K-M 模型在計算透明的介質上會有所限制，且不能反映出顏料混和後產生的化學變化 [9]。因此，若直接套用於半透明的水彩顏料混色預測，計算出來的混色會與實際結果所差異。

隨著類神經網路的發展，開始有人利用神經網路 (Neuron Network, NN) 來預測混色結果。如 Xu et al. [26] 的研究中利用 K-M 模型搭配淺層的神經網路來預測水墨畫的顏料混色，並得到卓越的成果。而近期 Chen et al. [7] 的研究使用了更深層的神經網路模型，與針對水彩混色提出的 Dataset，在水彩混色上也得到了相當好的成果。

本文平行於 Chen et al. 的研究，使用相同的水彩混色 Dataset，並提出一個不同的水彩混色預測模型。利用深度捲積神經網路的模型，並搭配文中提出的色彩感知誤差，來學習水彩反射光譜的混色模型。



本文提出的機器模型基於兩個主要的部份:

深度卷積神經網路: 近年來，深度卷積神經網路 (Convolutional Neuron Network, CNN) 的發展使圖形識別領域得到突破性的進步。而捲積神經網路也成功應用於其他領域，如：語音辨識等。本文作者希望透過深度卷積神經網路強大的特徵抽取能力，能夠對輸入的色彩特徵 (透射與反射) 中，抽出豐富且有用的資訊，以提升混色預測的準確度。

色彩感知損失: Chen et al. 的研究使用預測混色與真實混色 (Ground Truth) 在反射光譜上的誤差，作為神經網路訓練時所使用的損失函式 (Loss Function)。而本文使用了與前者不同的損失函式：將光譜轉換至色彩空間中的距離，作為損失函式。本文選擇的色彩空間是基於人類色感所得，因此透過色彩感知損失來訓練神經網路後，預期能在人類視覺上，得到更接近正確混色的結果。

透過以上兩部分結合所訓練出來的混色模型，比起傳統的方法在準確度上有大幅提升，預期在混色模擬與學習混色上會有更好的幫助。

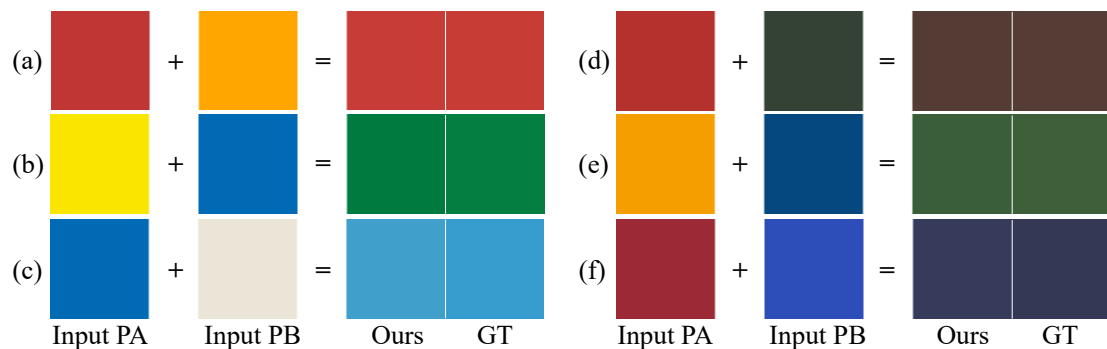


Figure 1.1: Part of color mixing results (a)~(f), showing different cases in our test set. With two input pigments PA and PB, our model can predict the results (Ours, left) which are almost indistinguishable to the ground truth (GT, right) of the real mixed pigment.



Chapter 2

Related Work

2.1 Kubelka-Munk Model for pigment color mixing

傳統的顏料混色的研究一般會選擇使用不同的物理模型或是經驗模型來計算顏料的混色結果，而在這些模型中，Kubelka-Munk [1931] 的模型最常被使用。

Kubelka-Munk 模型針對於均質且具等向性的薄層物質 (如顏料塗層)，利用兩個係數 K 與 S (分別代表吸收與散射) 之線性組合，來計算顏料混色的散射值與吸收值，進而計算混合顏料的反射值。因其模型簡易直觀，加上散射值與吸收值易於估算，許多數位繪畫相關的研究選擇 K-M 方法作為顏料混色模型。

較早的論文有 Haase and Meyer [1992]，其使用 K-M Model 計算出混和顏料的理論反射值，並利用 CIE XYZ 座標將反射光譜轉換為 RGB 色彩，顯示在電腦螢幕上。Curtis et al. [1997] [8] 的論文，提出了一套電腦上摹擬水彩畫的系統，其中的顏料模型使用了簡化的 K-M Model：使用三個 K, S 參數來代表 RGB 三原色，以兩個使用者選定的 RGB 顏色來計算出混色後的 RGB 值。Baxter et al. [2004] 的論文，提出一個互動式的數位繪畫系統，其顏料模型亦使用了 K-M Model 推導其混色函式：量測數種基本顏料的反射光譜，利用顏料反射數值，透過推導後的 KM model 來計算出混色結果的反射光譜轉換為 RGB 值，並使用了 GPU 來加速計算。

此後與顏料混色有關的研究，亦廣泛的使用 KM Model 來預估混色的結果。

然而，K-M 理論在實際應用中，其理論計算與具體實踐之間常出現差異，究其原因可歸納為四個因素。



1. K-M 理論假設所有介面的折射係數相同，不考慮介面引起的反射，然而在混色模擬中”顏料與空氣”和”顏料與紙”的介面必然存在，直接使用 KM Model 可能造成誤差。

2. K-M 理論是以不透明介質為前提提出來的，而水彩使用的顏料是透明性或半透明性的，在某些情況下吸收值接近零，使 KM Model 的預測與真實情況有所誤差。

3. 實際上在暗色調處(吸收力強的介質)，相當多的光線在散射前已被吸收，所以這些進入色層的光束不會擴散，散射值接近於零，致使 KM Model 的預測結果出現較大差異。

4. K-M 理論假設不同介質之間不會產生化學反應或是電性作用，而水彩混色可能產生上述的交互作用，使得介質中粒子的大小變得不均勻，使 KM Model 的預測與真實情況有所誤差。

雖然 KM Model 在先前的論文中都能得到不錯的摹擬效果，但是在物理上未必能完全符合水彩混色的，因此我們希望能找到更好的方法去建立水彩顏料的模型。

2.2 Neuron Network for Color Mixing

在人工神經網路 (Artificial Neuron Network, ANN) 發展的早期 (1990 年代)，開始有許多研究使用 ANN 來預測一些物理與化學現象，如 Bishop et al. [1991] [5] 提出以神經網路來完成曲線擬合的預測。而水彩混色可以視為：兩個顏料中與色彩相關的物理性質作為輸入，來對混出顏色的反射光譜來做曲線擬合。

Westland et al. [2002] [24] 的研究中首先提出利用 ANN 的方法，來建立顏料混色模型，並與 KM Model 做比較。但可能因當時 NN 的技術尚未成熟，其實驗結果顯示 ANN 未能勝過 KM Model。但該篇論文亦提出將 ANN 與 KM 模型結合的想法，並預期 ANN-KM 模型有機會勝過純粹的 KM Model。

而 Xu et al. [2007] 提出的水墨畫模擬系統，便使用了 ANN 模型搭配了 KM 方法，得到了相當好的結果。其所使用的顏料資料庫 (Dataset) 考量了 KM 模型會用到的吸收與散射 (K 與 S)，同時也使用了顏料的反射值與透射值 (R 與 T)，與兩顏料的份量，來作為 NN 的輸入。但為了簡化輸入的複雜度，因此只考慮了三個波長，分別對應純色的紅, 綠, 藍三原色。並加上基於 KM Model 的核函式 (kernel function) 作為輸入來幫助模型計算。其結果顯示出 ANN-KM 模型的混色結果與實際的混色已相當接近，並且優於單純使用 ANN 的方法。

隨著 ANN 在近期的快速發展，我們現在有更好的工具與更多的方法來訓練更複雜的深度神經網路 (Deep Neuron Network)。而 DNN 的進步在其他領域上已帶來許多突破，近期也在水彩混色模型上有新的進展。

Chen et al. [2018] 提出新的水彩混色模型，使用了資料庫與 DNN 模型。其顏料資料庫考量了可見光光譜上的反射與透射光譜，比起之前使用 ANN 的研究，大幅提升了顏料輸入特徵的維度。模型上並未加入 KM Model，而是使用了更深層的 DNN 來學習水彩顏料混色的模型，預測顏料混色的反射光譜，也得到了與實際混色相當接近的成果。

本論文平行於於 Chen et al. 研究，使用相同的水彩混色資料庫，但提出不同的方法來做出水彩混色模型。利用深度捲積神經網路 (Deep Convolutional Neuron Network, 後面以 CNN 簡稱) 來建立模型，並搭配文中提出與 RGB 色彩空間相關的色彩感知誤差，來學習水彩混色光譜的混色函式，以期能得到更加準確的結果。

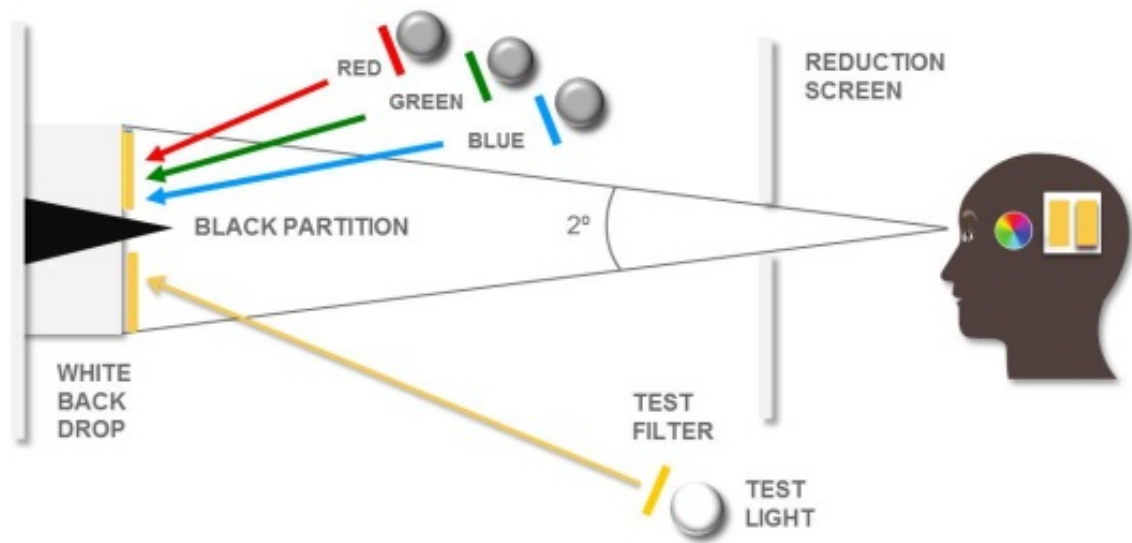


Figure 2.1: 圖為 Wright 與 Guild 的實驗設計，觀測者利用可控亮度的三原色來匹配測試顏色。

2.3 CIE Color Matching Functions

人的眼睛內有三種辨別顏色的錐形感光細胞，分別對應短 (S, 420-440nm)、中 (M, 530-540nm) 和長 (L, 560-580nm) 波長的光感受器，此三個波段分別對應黃綠色、綠色和藍紫色的光。雖然三種細胞並不是分別對紅色、綠色和藍色最敏感，但這三種光可以分別對三種錐形細胞產生刺激。因此，根據三種視錐細胞的刺激比例，便能描述任一種顏色的感覺。

色彩空間指的是用一種客觀的方式敘述顏色在人眼上的感覺，通常需要三色刺激值 (Tristimulus values)。首先先定義三種主要顏色 (primary color)，再利用顏色疊加模型，即可敘述各種顏色 (但三種主要顏色未必需要是真正的顏色)。

而在顏色感知的研究之中，**CIE 1931 XYZ** 色彩空間是第一次基於人眼對於色彩感知度量建立色彩空間的嘗試，它是幾乎所有其它色彩空間的基礎。在 1920 年代，分別由 Wright [25] 和 Guild [10] 獨立進行了一系列人類視覺實驗，提供了 CIE XYZ 色彩空間規範的基礎。

實驗設計如圖 2.1 所示，使用 2 度視角的圓形螢幕。投影螢幕的一半投上測試顏色，另一半投上觀察者可調整的顏色，中間用黑色物體做分隔。可調整的顏色是紅綠藍三原色的混合，三色有固定的色度，但觀察者可調整的顏色亮度。觀察者會調整三種原色光的亮度，直到觀察到混合的顏色與測試顏色匹配為止。然而，不是所有顏色都單純靠觀察者控制亮度來匹配。在前述方法無法匹配顏色的情況下，會將三原色中的其中一種投影到測試顏色上 (實際上是紅色)，用餘下兩種原色混合與其匹配。對於這種情況，疊加到測試顏色上原色的數量被算做是負值。通過這種方式，可以覆蓋完整的人類顏色感知。對應可見光波段，將每種原色的匹配時所使用的數量繪製為對測試顏色的波長的函數，這三個函數稱為此實驗的“色彩匹配函數”(Color Match Function)。(圖 2.2)

Wright 和 Guild 的實驗結果被總結為標準 CIE 1931 RGB 顏色匹配函數，之後為了符合一些標準，CIE 委員會確立出 CIE RGB 到 XYZ 空間的線性轉換，得出

標準的 CIE 1931 XYZ 色彩空間。反射光譜與 XYZ 色彩匹配函數積分後可以得到 XYZ 值，便可利用 XYZ 值轉換到欲使用的色彩空間。



本文利用以色彩匹配函數為基礎的感知誤差 (Perceptual Loss)，來幫助 CNN 模型在訓練時，學習如何做曲線擬合才能在 sRGB 色彩空間上能達到最小的誤差，進而提升混色結果在顯示器與人類視覺上的正確度。

CIE 1931 色彩空間的一個問題是它沒有給出估量色差的直接方式，在色度圖上距離無法均勻地對應兩個顏色之間的在感知上的差異度。

因此，直接基於了 CIE 1931 XYZ 色彩空間，CIE L*a*b* (CIELAB) 色彩空間被設計來測量色彩在感知上誤差。在 L*a*b* 模型中均勻改變對應於在感知顏色中的均勻改變，所以在 L*a*b* 空間中任何兩個顏色的相對感知差異。可以透過把每個顏色視為在 (L*a*b*) 三維空間中一個點，並計算在顏色點之間的歐幾里得距離。在 L*a*b* 色彩空間中的這個歐幾里得距離是 ΔE_{Lab} 。故本論文使用 ΔE_{Lab} 來評估預測混色與真實混色的色差。

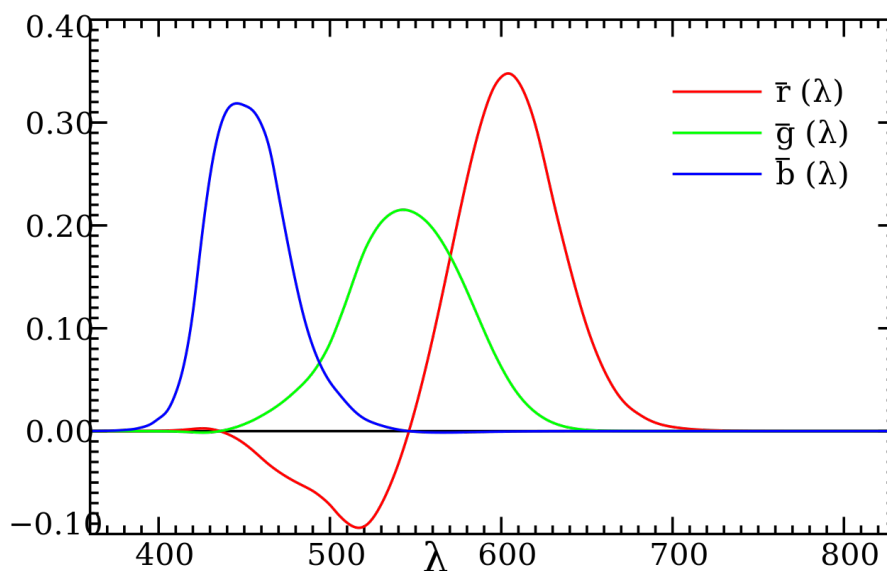


Figure 2.2: CIE 1931 RGB 顏色匹配函數。顏色匹配函數是匹配水平刻度標示的波長的單色測試顏色所需要的原色數量。

2.4 Development of Convolutional Neuron Network

自 LeCun 提出的 shared-weights network [17](後被稱作捲積神經網路, CNN) 後, 人工神經網路在圖形辨識上開始有所進展。在 1998 年, LeCun 所提出的 LeNet-5 [18], 使用在當時開創性的七層 CNN 架構, 來數間銀行用來辨識值票上的手寫數字(數位化成 32*32px 解析度)。然而, 受到當時電腦的計算能力限制, 尚無法處理更高解析度的圖片(需要更大且深層的 CNN)。

2005 開始有論文展現出 GPGPU 在機器學習上的價值 [22], 而後也有其他著作闡述如何應用 GPU 來訓練 CNN [6] [13] [4] [20]。直到 2012 年, AlexNet [15] 的提出在圖形辨識上有重大的突破, 其提出的深度捲積神經網路在 2012 ImageNet 物件分類競賽上得到了第一名的成績, 並且遠勝第二名的傳統架構。

AlexNet 之後, CNN 在電腦視覺方面成為了的新典範, 而後幾年也發展出了更新且更深的架構。深層的神經網路在結構上能夠擷取高中低層的特徵, 堆疊越多層的網路能擷取到更多不同層次的特徵。而近期的影像辨識方面的研究與競賽也應用了更深層的架構, 來取得更的結果, 如 VGG [21], GoogLeNet [23] 等等。

但更深層的網路在訓練的時候會遇到 degradation 的問題。Degradation 的問題是: 隨著神經網路深度加深, 通常準確率會提升至飽和狀態, 但再繼續加深會造成準確率快速下滑, 如圖 2.3 所示。這也說明更深層的神經網路的訓練會較為困難。

為了克服上述問題, He et al.[2015] [12] 提出了一種新架構, 稱為深度殘差網路 (Deep Residual Network)。透過修改網路, 加入殘差 (residual) 與自我映射 (Identity Mapping), 可以克服深層神經網路的訓練問題, 進而提升更深層神經網路的預測準確率。其團隊在 ImageNet 影像識別競賽上達到 3.57% 之錯誤率, 為 ILSVRC 2015 圖像分類競賽冠軍。

本研究發現深度殘差網路在顏料混色預測上也能達到相當好的精確度, 模型

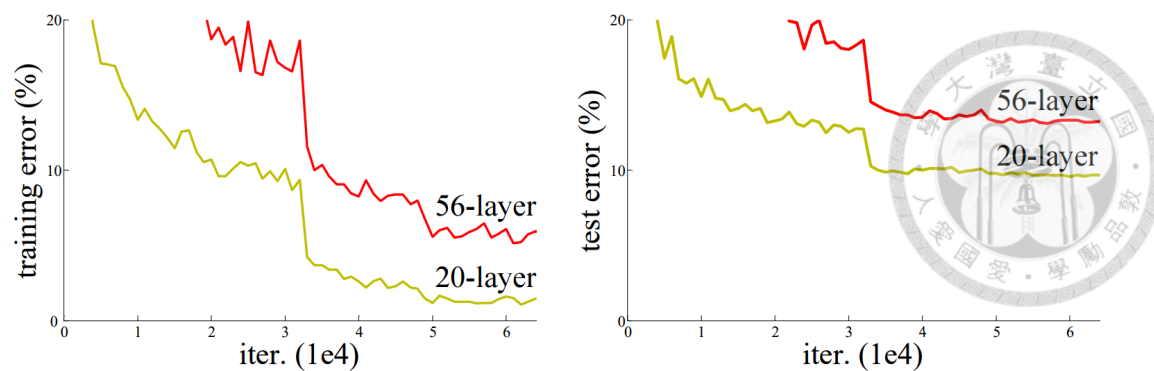


Figure 2.3: 圖例引用自 He et al. 之論文。顯示出無論在 training(左) 與 testing(右) 時，使用過深的”普通”神經網路會造成較高的誤差。

複雜度也低於較舊的架構 (VGG)，因此本研究最終採用深度殘差加上捲積神經網路，做為水彩混色的預測模型。



Chapter 3

Method

本文提出的混色模型分為兩個部分：前半部分利用深度卷積神經網路來抽取顏料輸入特徵，來預測混色的反射光譜。後半部分透過最小化色彩感知損失，讓預測出來的混色能在人眼觀測上接近實際混色。

如圖3.1所示，透過卷積神經網路強大的特徵抽取能力，可以對輸入的特徵(顏料之反射與透射光譜)抽出豐富且有用的資訊，以提升混色預測的準確度。可以把輸入的特徵看作是數張圖片，CNN 會從中抽取有用的局部特徵 (Local Feature)，而深層的神經網路則會讓特徵的層次更多，乃至全局特徵 (Global Feature) 也能在最後抽出來。本文採用的模型是修改過的 ResNet-18，其特性有助於深層模型的訓練，因此有助於抽取由淺至深層次的特徵。緊接著池化層 (Pooling layer)，將最後層的特徵壓縮以減少過擬合 (Overfitting)，並大幅降低模型的複雜度。最後的全連接層 (Fully Connected layer) 利用壓縮後的特徵，來預測最後混色的反射光譜。

後半段的損失函式部分，本文採用與 Chen et al. 使用曲線之間的數值誤差作為損失的不同的做法。本文提出使用基於人眼感知的色彩空間上的誤差 (Perceptual Loss) 作為損失函式，以期在混色預測結果上能在視覺上接近實際的混色結果。而實際上我們採用線性 RGB 色彩空間，因 RGB 為資訊世界中最常被使用的色彩空間，而 RGB 實際上也是以人類視覺為基礎的三色刺激值，因此能符合我們損失函式的要求。評估函式 (Evaluation Metrics) 我們則選用 LAB 色彩空間中的歐

式距離 ΔE_{Lab} ，為最常被使用的色差量測標準。

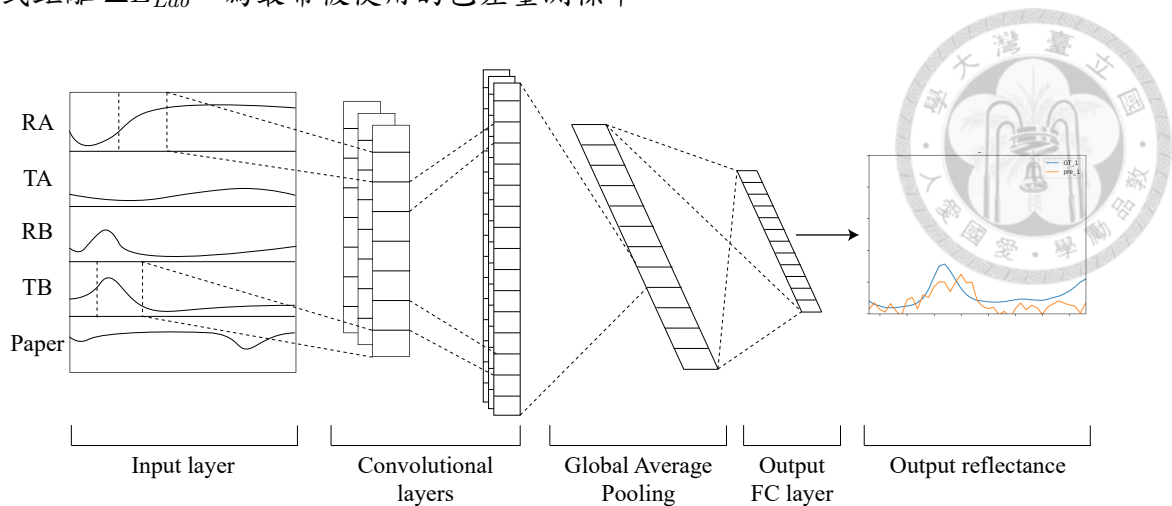


Figure 3.1: An illustration of the architecture of our model. We use a alternative ResNet18 to extract multi-level features from input reflectance and transmittance (RA, RB, TA, TB and Paper). A global average pooling layer is connected to the last conv. layer to reduce the number of parameters in our model. Fully connected layers after the pooling layer perform regression to minimize the perceptual loss between output and ground truth.

3.1 Convolutional Network for Pigment Mixing

利用捲積神經網路來抽取頻譜特徵的想法是啟發自 Abdel-Hamid et al.[2012] [2] 的研究。其研究利用 CNN 對一段語音的頻譜進行局部特徵抽取，來進行語音辨識。該研究說明了頻譜是可以透過 CNN 來抽取特徵，進行後續的預測。而我們的研究相似地也使用顏料的光譜做為輸入，因此能期待 CNN 從輸入光譜中抽取局部特徵來預測最終混色。

Inputs and Outputs

本文使用與 Chen et al. 相同的 Dataset，因此 input 為兩個水彩顏料的在可見光範圍 (380nm 至 780nm) 之反射與透射光譜 R1,R2,T1,T2，以及紙張的反射光譜 P。Training set 上共有 762 筆，testing set 上 261 筆。每個光譜在可見光範圍中採樣 41 個點。總共維度維 $41 * 5 = 205$ 維。與 Chen et al. 的方法略有不同的是，本文沒有將兩顏料分量加入 input 中。Output 為一樣 41 維的混色顏料反射光譜 R_{out} ，利用 color matching functions 可將 R_{out} 轉為 RGB 顏色。

Data Augmentation

作者在實作中發現 Chen et al. 的 Dataset 中，Input 顏色的順序並無做反轉。然而實際混色上即使 input 顏色反轉應該也有相同結果，故作者在 training 與 testing set 上做 input 顏色反轉，使得原本的 data 數量增加一倍。



Simple CNN experiment

作者首先嘗試簡單的 CNN 架構. 首先將 input resize 為 (41,5,1)，可以看作是大小為 41*5 的灰階圖片。緊接著堆疊兩層 Conv2d layer 抽取特徵，每層包含了 512 個 filter，並使用非線性的 Rectified Linear Unit 作為激勵函式 (activation function)。Kernel size 選擇使用 (9,1)，因為測試上發現此 size 能得到最佳的結果。在第二個 conv layer 後接一個 dropout 以避免 overfitting，亦能降低些許 error。conv block 後面接上三層 fully connected layer，size 分別為 512.256.41，最後一層為 output layer。透過 conv layer 所抽出的 feature 來預測輸出的混色光譜。雖然該方法以能得到視覺上相當接近 ground truth 的顏色，但該 model 的參數量過於龐大，在實際應用上可能會需要較大空間儲存模型。作者為改善參數量過多的問題曾經試過在 conv layer 輸出部分使用 average pooling 來做降維，但會使輸出的 error 提高。為了在不提升 error 的前提下又能降低模型參數量，作者發現採用了更深層的 conv net 可以解決此問題。

ResNet for color mixing

經實驗發現 ResNet18 在混色預測問題上得到比上述模型更好的結果，並且在參數量與深度上都比上述類似 VGG 的模型多。基本架構類似 He et al. 所提出的 ResNet18，但有小部分的修改 (如圖3.2所示)。其中，初始的 conv 層只使用 (1*1) 的 filter size，並且將接續的 MaxPooling 去除。第一和第二個 ResBlock(下圖藍色與綠色區域) 使用了 (9*1) 的 filter size，而後面的 ResBlock 使用 (3*1) 的 filter size。第一個 ResBlock 以後的 Block 輸入都會降維一次，其 input size 如下圖所示。在最後的 ResBlock 後方續接一個 Global Average Pooling，將最後的 feature map 降維以減少參數量與 overfitting。最後一層則為 fully connected layer，將 pooling 後的 feature 用來做混色顏料的預測。最後的輸出一樣是 41 維的混色 spectrum。

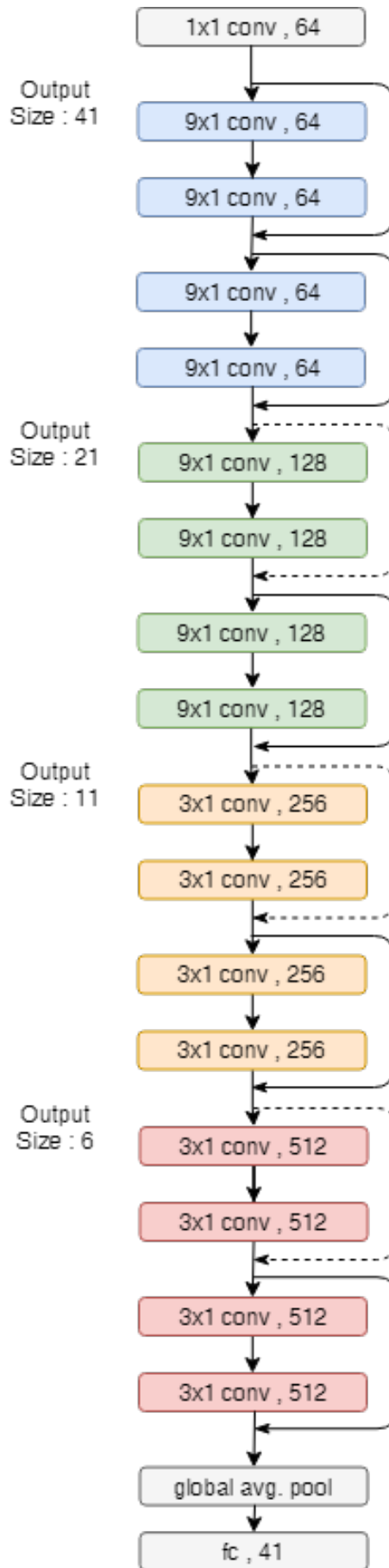


Figure 3.2: An illustration of our modified ResNet18 network, which is built with different filter sizes, pooling and output layers to generate a reflectance spectrum.

3.2 Perceptual Loss Function to Color

我們定義了一個基於色彩感知的誤差函式，來評估混色在色彩空間上的距離，用來降低混色結果與真實混色在人眼感知上的差異。然而，在 CNN 模型的最後面產生的是混色的反射光譜，我們需要經過一些轉換後才能轉換到色彩空間中。



Spectrum data to CIE XYZ

對於給定的顏料的反射光譜分佈 $S(\lambda)$ ，乘以參考光源的光譜功率分佈 $I(\lambda)$ 的 XYZ 三刺激值如下計：

$$\begin{aligned} X &= \frac{1}{N} \int_{\lambda} x(\lambda) S(\lambda) I(\lambda) d\lambda \\ Y &= \frac{1}{N} \int_{\lambda} y(\lambda) S(\lambda) I(\lambda) d\lambda \\ Z &= \frac{1}{N} \int_{\lambda} z(\lambda) S(\lambda) I(\lambda) d\lambda \end{aligned} \quad (3.1)$$

其中 $x(\lambda)$ ， $y(\lambda)$ 和 $z(\lambda)$ 是基於 CIE 1931 規範的標準觀察員函式 (CIE standard observer functions) (2 度) (如圖 3.3 所示)。積分是對可見光譜上的頻率 λ (約 380 至 780 nm) 計算的。

而上式中的 N 為參考光源的亮度，是由參考光源的光譜功率分佈 $I(\lambda)$ 乘上代表光譜亮度的觀察者函數 $y(\lambda)$ ，對可見光頻率範圍做積分，如下式。其目的為標準化 XYZ 的數值，為色彩空間轉換做準備。

$$N = \int_{\lambda} y(\lambda) I(\lambda) d\lambda \quad (3.2)$$

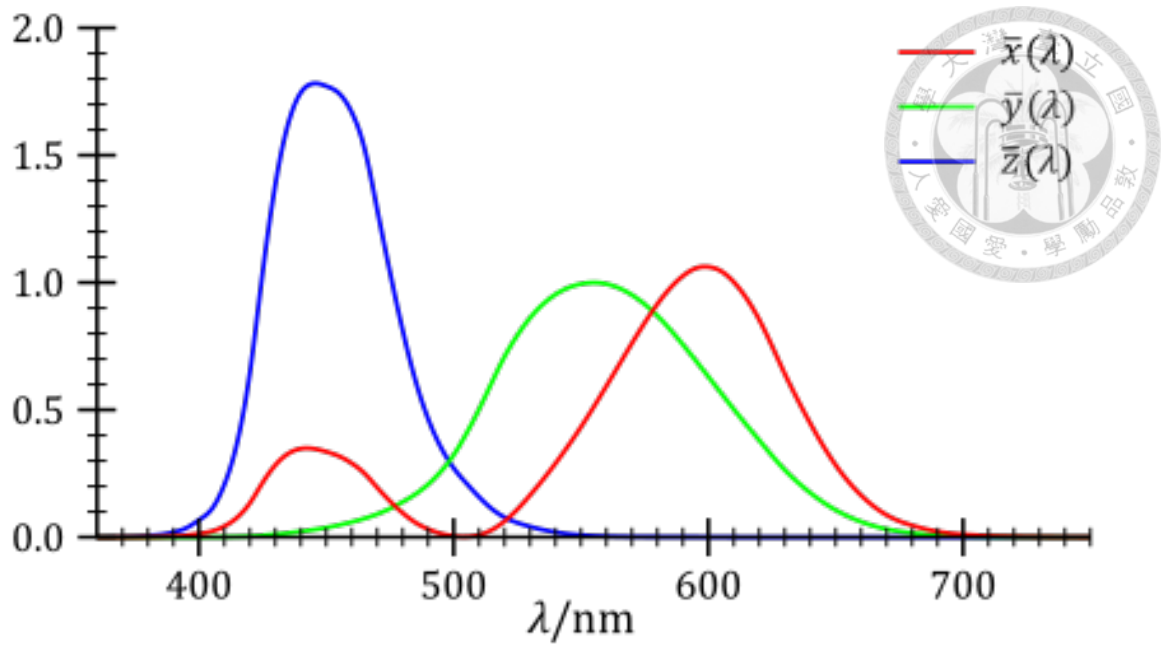


Figure 3.3: 圖為在 380nm 至 780nm 之間的 CIE XYZ 標準觀察的色彩匹配函數。

然而，積分中的函數(式子 (3.1))，只存在於理論之中。實際計算上，是以離散取樣的形式存在，所以積分被替換為求和：

$$\begin{aligned}
 X &= \frac{1}{N} \sum_i x_i S_i I_i \Delta\lambda \\
 Y &= \frac{1}{N} \sum_i y_i S_i I_i \Delta\lambda \\
 Z &= \frac{1}{N} \sum_i z_i S_i I_i \Delta\lambda \\
 N &= \sum_i y_i I_i \Delta\lambda
 \end{aligned}
 \tag{3.3}$$

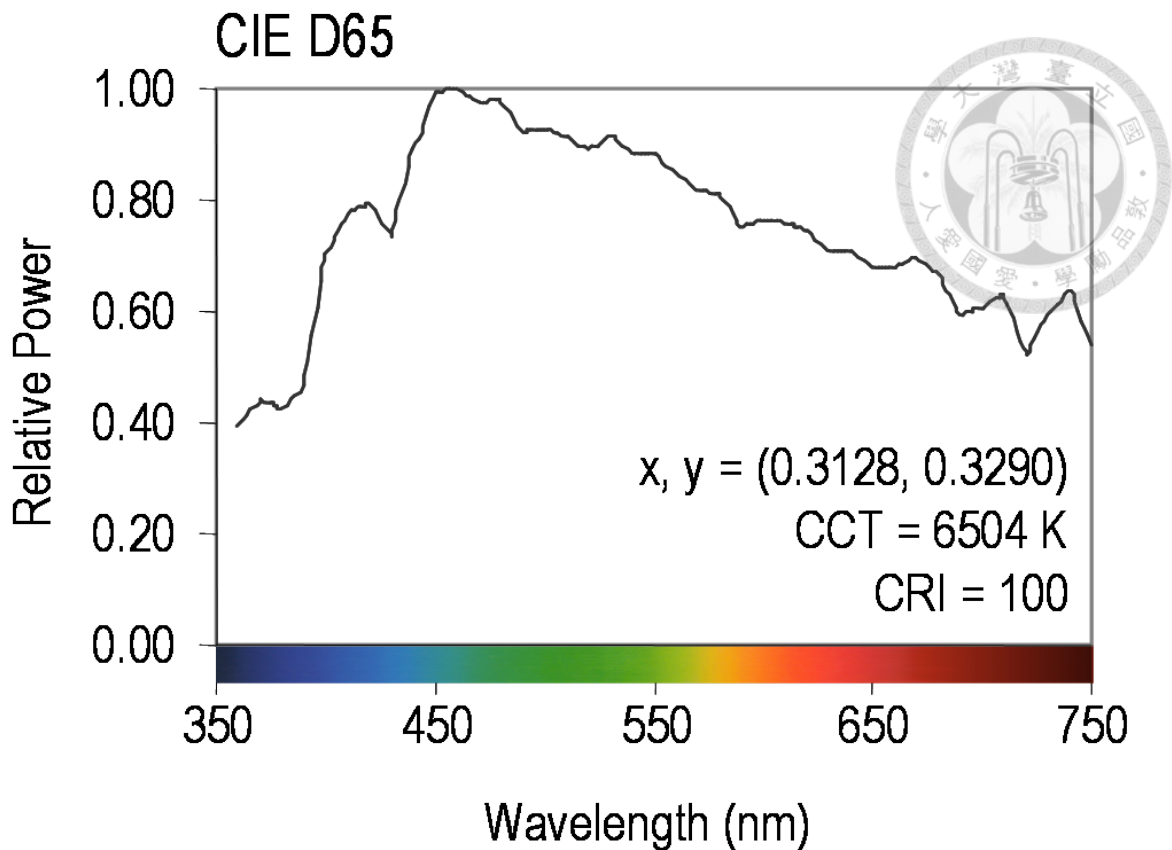


Figure 3.4: Spectral power distribution of D65.

Standard Illuminant

使用參考光源 (Standard Illuminant) 的原因是：如果考慮在黑暗的房間中觀察反射樣本 (顏料) 時，是看不到任何物體的。因為樣本不會發光，唯一可以看到的方式是先照亮它。很明顯，樣本顏色的受其接收的照明類型的影響。為了使顏色樣本的 XYZ 值明確，光源必須以某種方式識別並與之關聯。由於存在無數可能的光源，因此這帶來了問題。因此，為了統一標準，而是使用參考光源。常見的參考光源是”D50”，”D65”等。本文使用的是”D65”的參考光源，圖3.4為 D65 光源之標準光譜功率分佈。概略上，D65 光源接近一般的日光。

CIE XYZ to Linear RGB Euclidean Distance

給定均一化 XYZ 色彩值 (其數值範圍為 [0.0, 1.0])，若其參考光源與欲轉換至的 RGB 系統相同，則 XYZ 轉換至線性 RGB 空間只需要一個步驟，如下式。



$$\begin{bmatrix} r \\ g \\ b \end{bmatrix} = \begin{bmatrix} M \end{bmatrix}^{-1} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (3.4)$$

欲自 XYZ 空間轉換至線性 RGB 空間 (使用 D65 參考光源)，則上式的矩陣 $[M]^{-1}$ 使用下列數值。

$$\begin{bmatrix} M \end{bmatrix}^{-1} = \begin{bmatrix} 3.240479 & -1.537150 & -0.498535 \\ -0.969256 & 1.875992 & 0.041556 \\ 0.055648 & -0.204043 & 1.057311 \end{bmatrix} \quad (3.5)$$

轉換至 RGB 色彩空間後，我們選擇用 RGB 色彩空間上兩顏色點的歐式距離，來評估預測的混色與真實混色的距離。

$$\Delta E_{rgb} = \sqrt{(\Delta r)^2 + (\Delta g)^2 + (\Delta b)^2} \quad (3.6)$$

在神經網路的訓練上，我們對於整個訓練集合 (Training set) 上的所有結果 ΔE_{rgb} 的取平均，做為訓練神經網路所使用的損失函式，如下式。其中 N 為訓練集中的樣本數量，i 代表訓練集合中第 i 個樣本。

$$\Delta Loss_{rgb} = \frac{1}{N} \sum_i^N \sqrt{(\hat{r}_i - r_i)^2 + (\hat{g}_i - g_i)^2 + (\hat{b}_i - b_i)^2} \quad (3.7)$$

以上轉換流程皆可微分 (Differentiable)，因此可以在前述的模型訓練中作為損失函式使用。

Color Difference Metrics : CIE LAB Delta E function

由於 RGB 色彩空間並不是一個對於感知均勻的空間，因此對於顏色感知的差距計算實際上並不精準。因此我們在衡量模型好壞時，選擇 CIE 1976 $L^*a^*b^*$ 色彩空間 (俗稱 CIELAB)。



CIELAB 得自 CIEXYZ 色彩空間，在感知均勻性上優於 XYZ 以及 RGB 色彩空間。 $L^*a^*b^*$ 色彩空間的數值，以三維的正交基底表示： L^* 值密切匹配人類亮度感知， a^* 值對應紅色和綠色之間的位置，和 b^* 值對應黃色和藍色之間的位置。

$$\begin{aligned}L^* &= 116f\left(\frac{Y}{Y_n}\right) - 16 \\a^* &= 500f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right) \\b^* &= 200f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\end{aligned}\tag{3.8}$$

其中，

$$\begin{aligned}f(t) &= t^{1/3} && \text{if } t > (24/116)^3 \\f(t) &= (841/108)t + 16/116 && \text{if } t \leq (24/116)^3\end{aligned}\tag{3.9}$$

以及，

$$X_n = 0.95047, \quad Y_n = 1.00, \quad Z_n = 1.08883\tag{3.10}$$

X,Y,Z 代表樣本的轉換至 XYZ 空間的三色刺激值， X_n, Y_n, Z_n 代表色彩空間所使用參考照明的白點。我們使用 D65 標準光源，白點 XYZ 值如式 (3.10) 所示。數值同樣經過歸一化，故 Y_n 值為 1。

在 $L^*a^*b^*$ 空間中均勻改變對應於在感知顏色中的均勻改變。所以在 $L^*a^*b^*$ 空間中任何兩個顏色的相對感知差別，可以通過把每個顏色處理為 (L^*, a^*, b^*) 的三維空間中一個點，並計算在兩點之間的歐式距離。在 $L^*a^*b^*$ 空間中的這個歐式距離為 ΔE_{Lab} 。



$$\Delta E_{*ab} = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2} \quad (3.11)$$

其中，

$$\begin{aligned} \Delta L^* &= L_1^* - L_2^* \\ \Delta a^* &= a_1^* - a_2^* \\ \Delta b^* &= b_1^* - b_2^* \end{aligned} \quad (3.12)$$

對於一般的觀察者來說：

- $\Delta E_{Lab} < 2$ 時，觀察者幾乎無法察覺到色差。
- $3.5 < \Delta E_{Lab} < 5$ 時，在有些情況下觀察者可以察覺色差。
- $5 < \Delta E_{Lab}$ 時，觀察者能清楚察覺到顏色不同。

這些數值來自 Mokrzycki et al.[2011] [19] 的研究，雖然在少數情況下有誤差，但基本上符合我們的模型預測的結果。

以上轉換流程不可微分，無法在前述的模型訓練中作為損失函式使用，單純用於評估模型好壞。



Chapter 4

Experiments

4.1 Training Details

本文所使用的 model 與 loss function 都實作 Python 語言之上，並大量使用 TensorFlow 上與 Keras 軟體庫。

TensorFlow [1] 是一個開源軟體庫，由 Google Brain 團隊開發，用於實作各種機器學習的專案。Keras 則是一個開源的神經網路軟體庫，由 python 寫成。能夠執行在 Tensorflow 上，並且提供一個友善, 快速且具延展性的測試環境，來實作深度神經網路的模型。

訓練演算法使用隨機梯度下降法 (Stochastic Gradient Descent,SGD) 之中的 ADAM Optimizer [14] 來做模型權重優化，optimizer 參數使用 ADAM 的預設參數，loss function 選用前一張所提出的 RGB Perceptual Loss。訓練途中，若 500 次未能降低 loss 則將學習率 (learning rate) 減半，若一連 5000 次 epoch 都無法降低 loss 則提早停止訓練 (Early Stopping)。這些參數選擇基於我們實驗中的發現：太長的訓練次數並不會幫助模型進一步優化。

最大訓練 epoch 次數為 10 萬次，batch size 為 200。Model 選取採用 training loss 判斷，當前 epoch 的 training loss 比已存取的 Model 更小時，則存取權重，此方法可以確保訓練過程中能存取到 loss 最小的 model。

硬體使用 Nvidia GTX1080Ti 顯示卡，該卡擁有 11G 的顯示記憶體，3584 個 CUDA 計算核心 (1480Mhz)。電腦搭配 8 核心 16 線程之 CPU(AMD Ryzen 1700, 3Ghz)，以及 32G 記憶體。訓練次數通常會在 10k 次左右停止，費時約為 3.5 小時。



下圖為不同訓練次數的收斂曲線圖。下兩圖分別為訓練途中 RGB Loss 與 Delta E 之收斂曲線。可以注意到兩個點:

1. 無論是 RGB Loss 與 Delta E 的收斂情況都相當類似：這證明同為基於 XYZ 色彩空間的三色刺激值空間，在收斂情況上也會非常相似。
2. 訓練次數長對結果影響有限：訓練次數常雖然代表有更多次的機會可以挑選最佳模型，但本文所用的 ResNet18 模型在訓練後期會穩定收斂到一定值，再增加訓練次數對降低 loss 幫助有限。

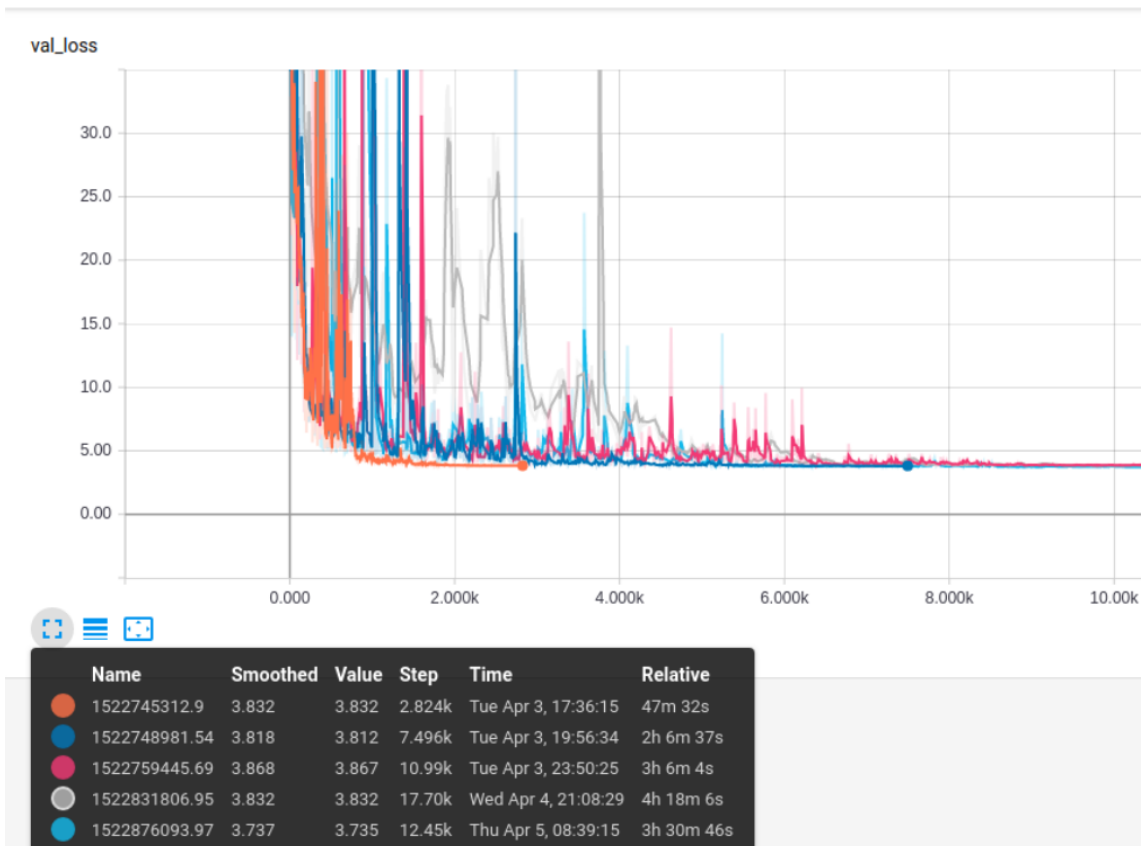


Figure 4.1: 圖為 RGB loss 收斂曲線，可以看到損失收斂曲線到了 8k 次訓練後已經到達飽和狀態。

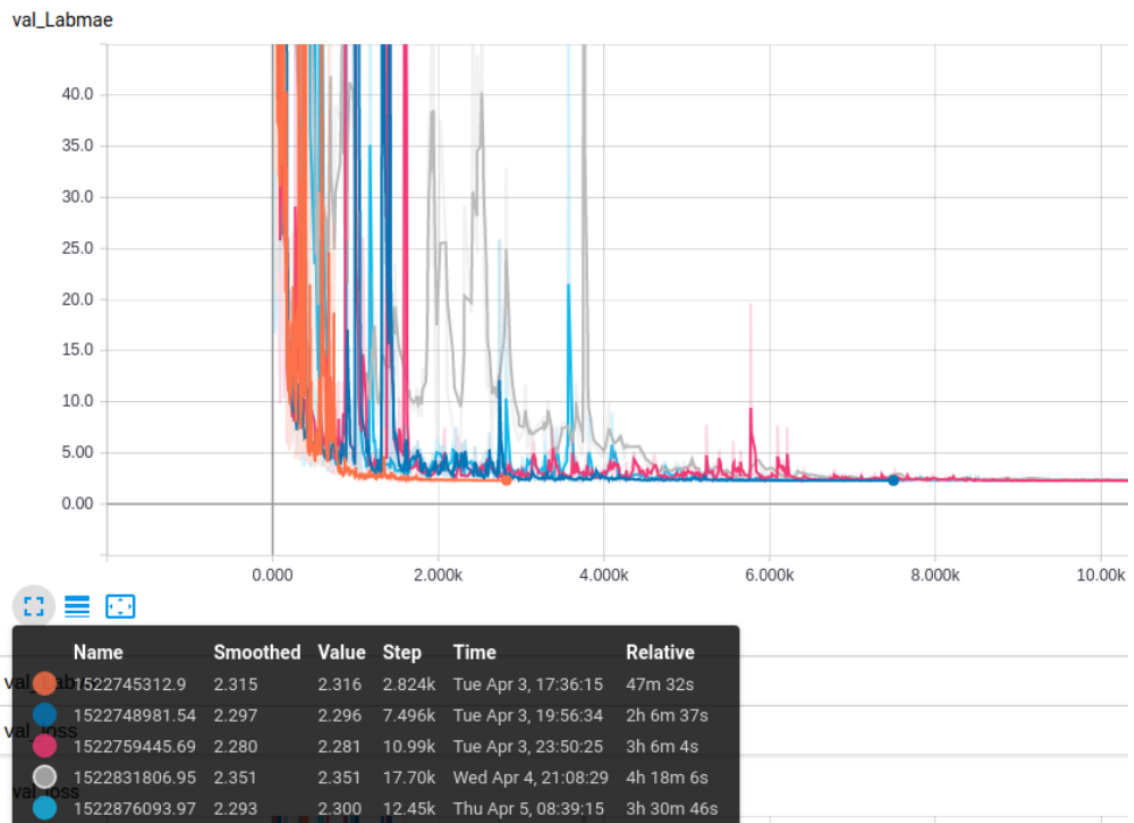


Figure 4.2: 圖為 Delta E 收斂曲線，與同為基於 XYZ 空間的 RGB loss 有類似的收斂曲線。

4.2 Result

統計結果

我們的模型在測試集上預測出來的混色在 ΔE_{Lab} 上可以達到平均 2.29，誤差極大值為 10.79。平均來說本文的結果已經接近了不易察覺色差的程度。

另外，在測試集上有 88.7% 的預測結果能達到 $\Delta E_{Lab} < 5$ ，有 79.7% 的預測結果能達到 $\Delta E_{Lab} < 3.5$ 。

```
Lab Delta E Euclidean Distance : 2.293936
Lab Delta E Max Value : 10.786460
Lab Delta E >3.5 Likelihood : 0.202869
Lab Delta E >5 Likelihood : 0.112705
```

Figure 4.3: 圖為訓練後的模型在測試集合上的 ΔE_{Lab} 統計結果。

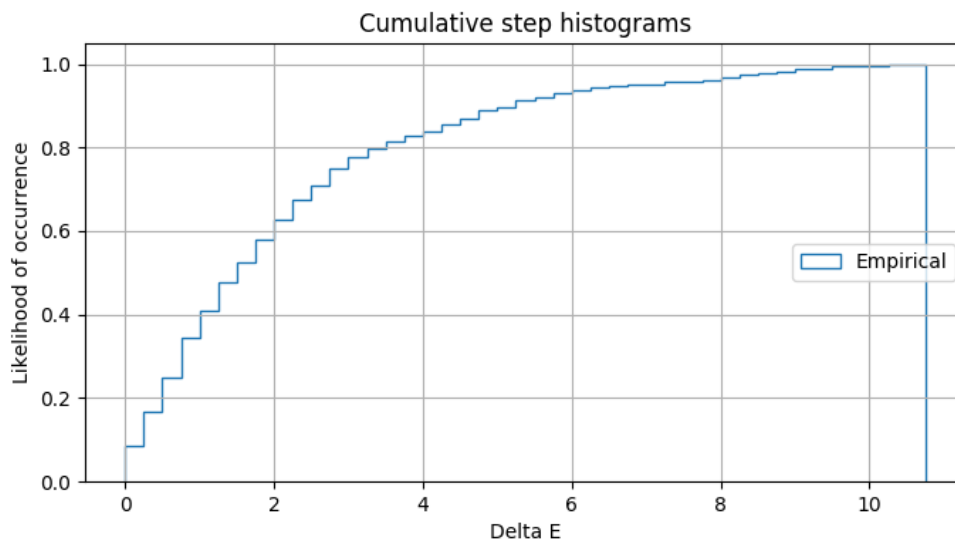


Figure 4.4: 圖為訓練後的模型在測試集合上的 ΔE_{Lab} 的累積分布圖。

色彩比較

下列數張圖為測試集中部分結果，並分成不同 ΔE_{Lab} 的區間說明。如上一章所述， $\Delta E_{Lab} < 3.5$ 代表預測結果與真實混色之間有可察覺的色差， $\Delta E_{Lab} < 5$ 代表有明顯色差。



Delta E < 2 : 下圖從左至右之 Delta E 為 0.58 , 1.0 , 1.93

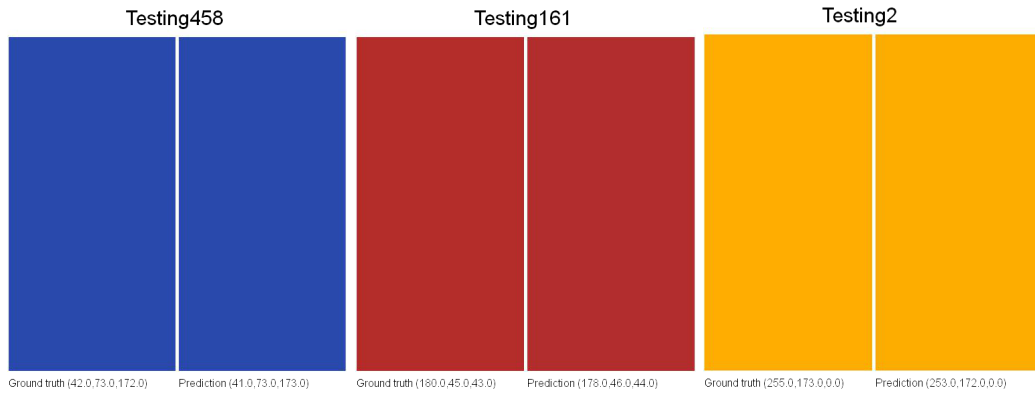


Figure 4.5: Delta E < 2 之部分結果

Delta E < 3.5 : 下圖從左至右之 Delta E 為 2.38 , 3.02 , 3.37

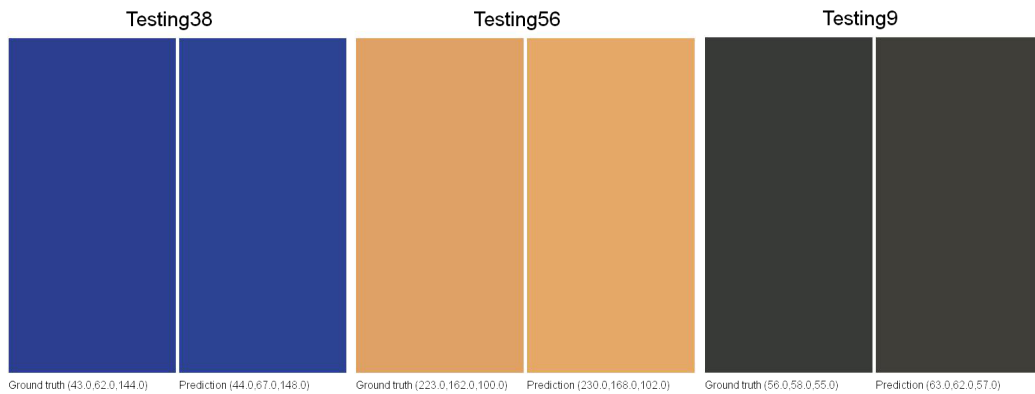


Figure 4.6: Delta E < 3.5 之部分結果

Delta E < 5 : 下圖從左至右之 Delta E 為 3.59 , 4.33 , 4.72

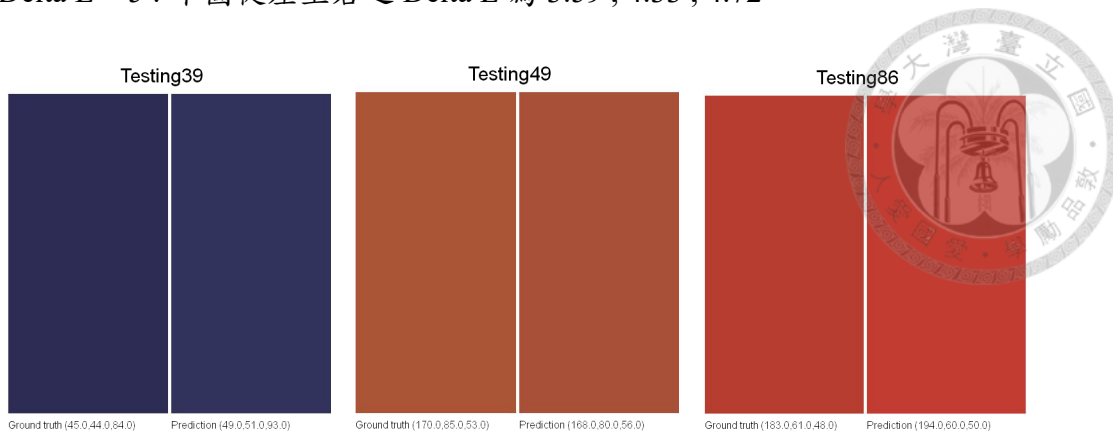


Figure 4.7: Delta E < 5 之部分結果

5 < Delta E : 下圖從左至右之 Delta E 為 5.49 , 5.98 , 6.64

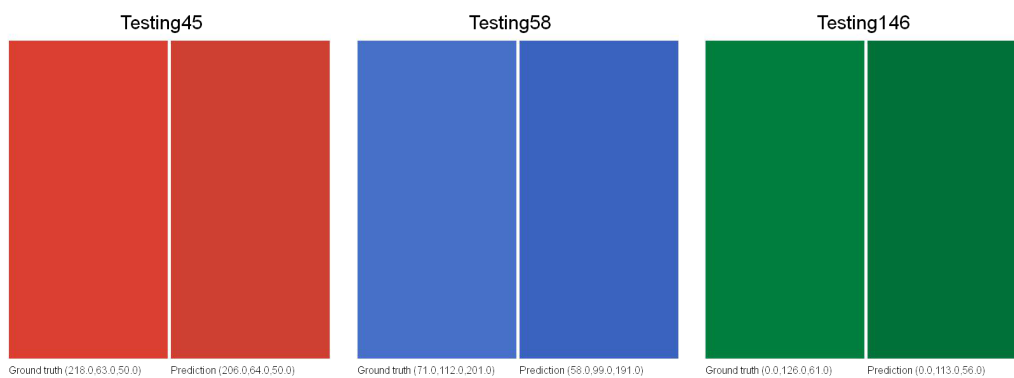


Figure 4.8: Delta E > 5 之部分結果

下圖從左至右之 Delta E 為 9.53 , 10.79 (為測試組中誤差最大者)

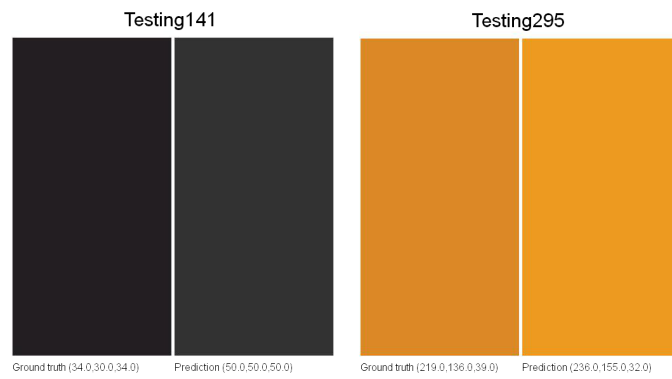


Figure 4.9: 測試集中之 worst case



RGB 顏色上的結果如下圖4.10所示，其中每個小色塊的左邊為 Ground Truth，右邊為預測出來的混色結果。結果由左至右，由上至下，依 ΔE_{Lab} 數值由小至大排列。由下往上數起第三排開始為 $\Delta E_{Lab} > 5$ 之結果，由下往上數起第五排開始則為 $\Delta E_{Lab} > 3.5$ 之結果。

可以看到上半部大部分的結果都與真實混色相當接近，甚至是無法分辨兩色差異，說明本文所提出的混色模型確實是在視覺上接近真實混色。

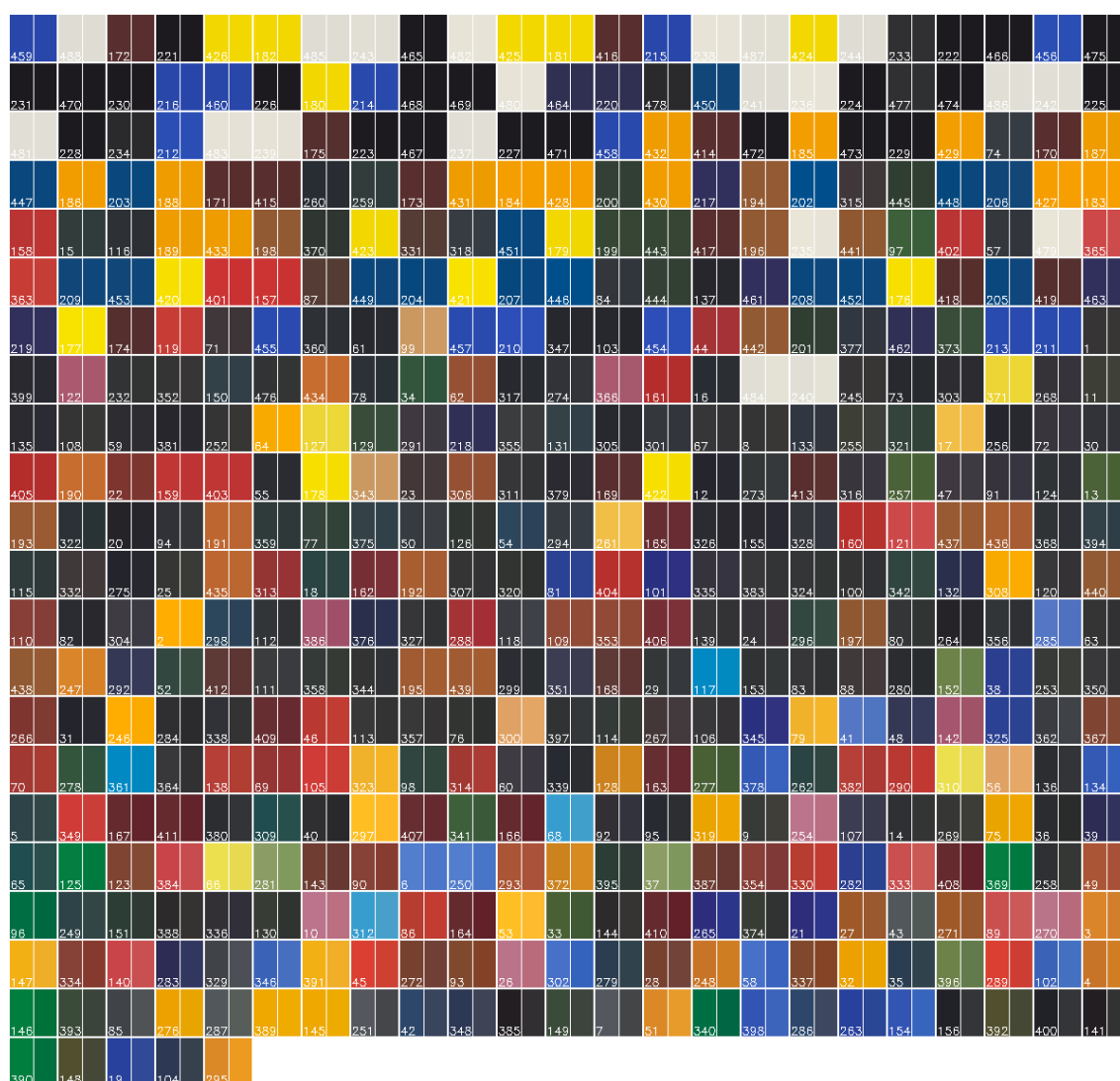


Figure 4.10: 圖為訓練後的模型在測試集合上的 RGB 顏色上的結果。其中每個小色塊的左邊為 Ground Truth，右邊為預測出來的混色結果。結果由左至右，由上至下，依 ΔE_{Lab} 數值由小至大排列。





Chapter 5

Discussion

Limitation in Dark Colors

雖然在色彩上，大部分情況下貼近真實混色。但本文所提出的方法是預測混色的反射光譜，所以在低反射度的混色上容易受到一些波動 (Noise) 的影響，而產生或多或少的色差。低反射度的反射光譜由於數值偏低，稍有偏移就會在色彩上顯得較亮，如下圖5.1所示。

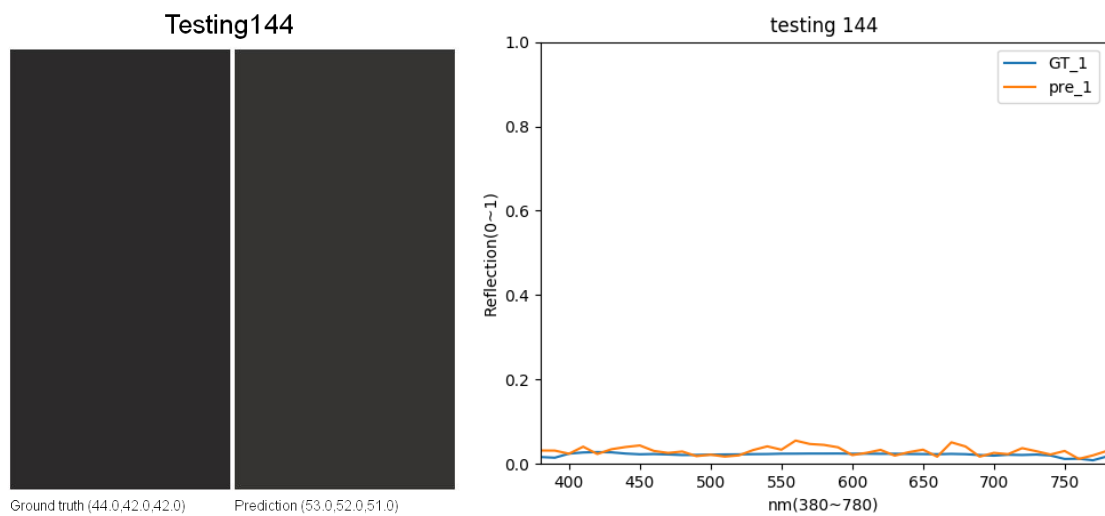


Figure 5.1: 此圖說明：暗色調混色容易因反射光譜的雜訊與飄移而造成色差。左圖左邊為實際混色，右邊為預測結果。右圖藍色線為實際混色光譜，橘色為預測結果。

New discovery on Predicted Reflectance Spectrum

本研究中另一個發現是：在預測的反射光譜上有波動的现象產生。

如下圖5.2所示，我們發現預測光譜會有抖動的現象。但即使有這樣大幅地抖動，不會影響在色彩空間中的準確度。從預測光譜可以發現，我們的方法所產生的光譜會偏重於700nm以下的區域，且在局部平均上與實際混色光譜類似。可能是我們的模型會使用色彩匹配的方式來在光譜上混色，如下圖場合即為：藍色波段加上少許綠色波段。

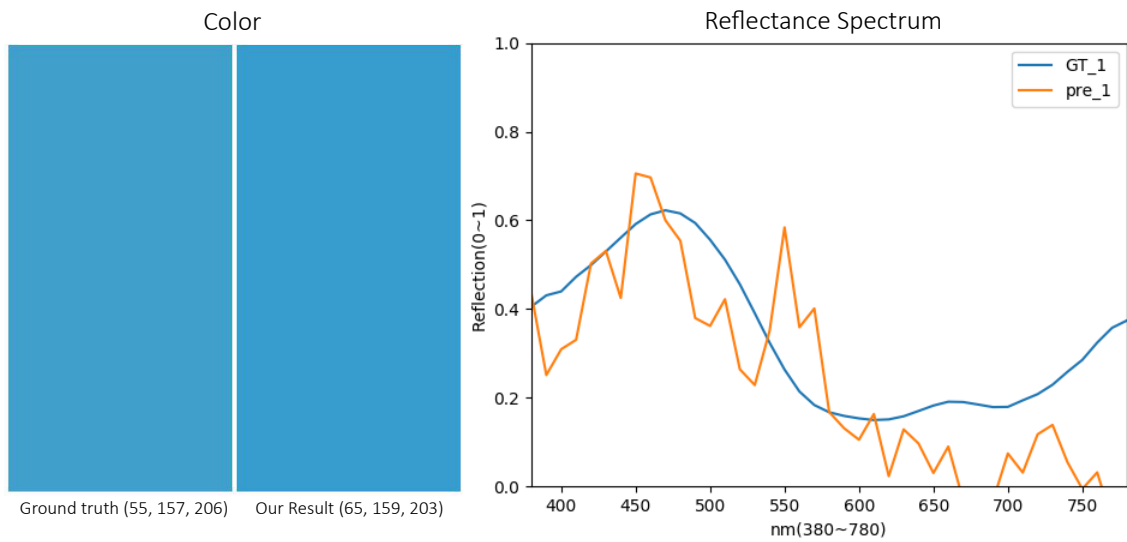


Figure 5.2: An illustration of the output reflectance of our model. The curve between visible wavelengths (380 to 750 nm) mainly focus on the short-term average, while the curve beyond 750nm can be totally ignored (being invisible to human eye)

而下圖5.3所示，可以從以下白色預測光譜範例觀察出：本文提出的模型在光譜上的偏好，主要落在 550-650nm 與 400-500nm 區間。同樣的，即便在光譜上兩曲線差異甚大，不會影響在色彩空間中的準確度。

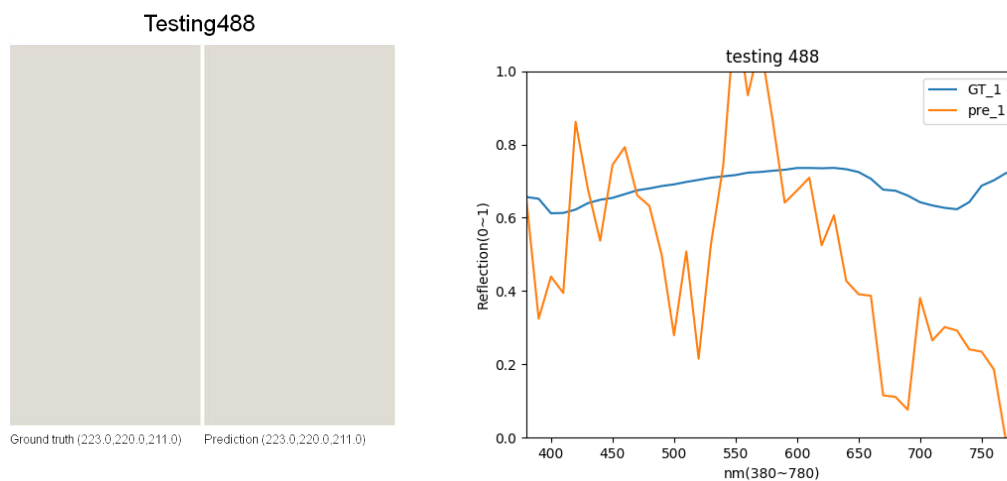


Figure 5.3: 此圖說明：白色混色在光譜上的偏重於三個區域。左圖左邊為實際混色，右邊為預測結果。右圖藍色線為實際混色光譜，橘色為預測結果。

Future Works

Zeiler et al. [2014] [27] 曾經透過特徵視覺化來幫助人們理解 AlexNet。若能將模型中的參數視覺化，或許能有助於我們理解上述的現象發生的原因。

改進混色模型是一個方向，若我們對模型的理解更加深入，或許能作出更好的混色模型。

應用方面，本文所提出水彩混色模型可以使用於未來的水彩繪畫模擬系統，幫助藝術家更好的製作數位水彩畫。另一個方面的應用可以利用混色預測，來製作一個混色匹配系統，給定 RGB 色彩來計算可能的水彩顏料混色配方。



Chapter 6

Conclusion

本論文所提出的卷積神經網路 (CNN) 的模型，應用於水彩混色預測，搭配上色彩感知誤差的損失函式來訓練模型，確實在平均上能達到相當接近真實混色的結果。此模型透過色感誤差的損失函式訓練後所產生的結果，雖然會有中幅度的波動，無法貼近真實頻譜，但在色彩上依然能與真實混色貼近。在暗色情況下可能會產生較大色差，是未來研究可以改進的點。

本研究提出的新方法及其帶來的成果，預期能為未來對於格種顏料混色的研究與數位繪畫模擬有所貢獻。

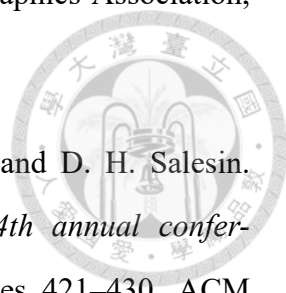


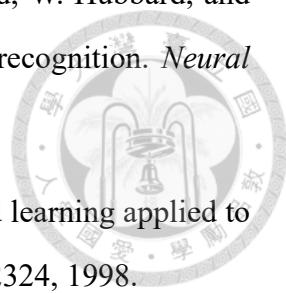


Bibliography

- [1] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al. Tensorflow: A system for large-scale machine learning. In *OSDI*, volume 16, pages 265–283, 2016.
- [2] O. Abdel-Hamid, A.-r. Mohamed, H. Jiang, L. Deng, G. Penn, and D. Yu. Convolutional neural networks for speech recognition. *IEEE/ACM Transactions on audio, speech, and language processing*, 22(10):1533–1545, 2014.
- [3] W. Baxter, J. Wendt, and M. C. Lin. Impasto: a realistic, interactive model for paint. In *Proc. of the 3rd international symposium on Non-photorealistic animation and rendering*, pages 45–148. ACM, 2004.
- [4] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle. Greedy layer-wise training of deep networks. In *Advances in neural information processing systems*, pages 153–160, 2007.
- [5] C. M. Bishop and C. Roach. Fast curve fitting using neural networks. *Review of scientific instruments*, 63(10):4450–4456, 1992.
- [6] K. Chellapilla, S. Puri, and P. Simard. High performance convolutional neural networks for document processing. In *Tenth International Workshop on Frontiers in Handwriting Recognition*. Suvisoft, 2006.
- [7] M.-Y. Chen, C.-S. Yang, and M. Ouhyoung. A smart palette for helping novice painters to mix physical watercolor pigments. In E. Jain and J. Kosinka, editors, *Proc.*

of *EuroGraphics 2018, Posters*, page April 2018. The Eurographics Association, 2018.

- 
- [8] C. J. Curtis, S. E. Anderson, J. E. Seims, K. W. Fleischer, and D. H. Salesin. Computer-generated watercolor. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 421–430. ACM Press/Addison-Wesley Publishing Co., 1997.
- [9] P. Edström. Examination of the revised kubelka-munk theory: considerations of modeling strategies. *JOSA A*, 24(2):548–556, 2007.
- [10] J. Guild et al. The colorimetric properties of the spectrum. *Phil. Trans. R. Soc. Lond. A*, 230(681-693):149–187, 1931.
- [11] C. S. Haase and G. W. Meyer. Modeling pigmented materials for realistic image synthesis. *ACM Transactions on Graphics (TOG)*, 11(4):305–335, 1992.
- [12] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proc. of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [13] G. E. Hinton, S. Osindero, and Y.-W. Teh. A fast learning algorithm for deep belief nets. *Neural computation*, 18(7):1527–1554, 2006.
- [14] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [15] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
- [16] P. Kubelka. New contributions to the optics of intensely light-scattering materials. part i. *Josa*, 38(5):448–457, 1948.

- 
- [17] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4):541–551, 1989.
- [18] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [19] W. Mokrzycki and M. Tatol. Colour difference ΔE -a survey. *MGV*, 20(4):383–411, 2011.
- [20] C. Poultney, S. Chopra, Y. L. Cun, et al. Efficient learning of sparse representations with an energy-based model. In *Advances in neural information processing systems*, pages 1137–1144, 2007.
- [21] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [22] D. Steinkraus, I. Buck, and P. Simard. Using gpus for machine learning algorithms. In *Document Analysis and Recognition, 2005. Proceedings. Eighth International Conference on*, pages 1115–1120. IEEE, 2005.
- [23] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, et al. Going deeper with convolutions. *Cvpr*, 2015.
- [24] S. Westland, L. Iovine, and J. M. Bishop. Kubelka-munk or neural networks for computer colorant formulation? In *9th Congress of the International Colour Association*, volume 4421, pages 745–749. International Society for Optics and Photonics, 2002.
- [25] W. D. Wright. A re-determination of the trichromatic coefficients of the spectral colours. *Transactions of the Optical Society*, 30(4):141, 1929.
- [26] S. Xu, H. Tan, X. Jiao, F. Lau, and Y. Pan. A generic pigment model for digital painting. In *Computer Graphics Forum*, volume 26, pages 609–618. Wiley Online Library, 2007.

- [27] M. D. Zeiler and R. Fergus. Visualizing and understanding convolutional networks.
In *European conference on computer vision*, pages 818–833. Springer, 2014.

