

# 國立中正大學資訊工程研究所

## 碩士論文

醫學影像專題一

Medical Image Project1

張世亞、蔡旻勳

指導教授：林維暘

中華民國 109 年 4 月

# 誌謝辭

在這門課的兩個月裡，首先要感謝也要對不起我們的指導教授林維暘，我們不是一位非常優秀、十分用功的學生，是老師不厭其煩的一再細心教導與指引方向，令愚笨學生的我們能夠有所成長，才使得本論文得以順利完成，在此謹敬上最高謝意與感激

其次，感謝我們同窗的研究夥伴——軒綸、耘愷和在學期間所有同學、學長姐、學弟妹，有你們與我們共度這兩個月時光讓我們研究生生活更加美好充實與豐富。還要感謝我們大學的好友以及學弟妹們，在這孤單期間有你們陪伴使我們的生活更多彩多姿也讓我們充滿了開心。再來感謝在嘉義的好朋友與好哥們，雖然這段期間並不常見面，但是有你們的關心與打氣，讓我們更有動力能在繁忙的課業中順利撐過來。

另外特別感謝幾位我最重要的貴人——昀倫、宏瑜、佑安、麥克，當我們心情陷入低潮時是你們給予我們支持與鼓勵，讓我們學會成長、讓我們充滿勇氣、讓我們更堅強更勇敢的去面對一切困境，沒有你們也就沒有今天堅強的我們，謝謝你們，你們是我們一生中永遠最棒的好朋友。

最後，我們要感謝我們的父母、家人以及我們最敬愛的上帝，有你們細心栽培與照顧，讓我們無後顧之憂的學習與完成學業，所以我們的成就也是你們的驕傲。謹以此篇論文獻給我們的師長、家人、同學以及所有關愛我們的朋友，謝謝你們。

## 目錄

摘要.....	4
介紹.....	5
相關背景知識.....	5
方法.....	5
實驗.....	7
結論.....	11
參考.....	11

# 摘要

本篇論文目的在於使用深度學習針對醫學影像之息肉狀脈絡膜血管病變 (PCV)[1]，脈絡膜新生血管 (CNV)[2] 進行辨識與分類。息肉狀脈絡膜血管病變，為老年性黃斑部病變亞種之一，會造成視網膜黃斑部出血、積水，是目前老人視力下降的最重要原因之脈絡膜新生血管則同樣為老年性黃斑部病變亞種之一，但所造成的脈絡膜新生血管引發的視網膜出血的機制不同。由於目前兩者尚未出現公認的獨特特徵基因，導致人們無法透過分析基因關聯性來辨別兩者，不力於早期的診斷與治療[3]。目前常見的分辨方法為透過豐富經驗的醫生對兩者之間的細部差異(例如：血管半徑，異常分支血管網等等)進行辨識，既消耗人力又耗時。我們本次實驗中分別採用了 VGG16, ResNet50, ResNet101, ResNet152 四種架構加上遷移式學習並搭配不同優化器與進行資料擴增進行試驗，學習上述兩者影像之間的細微差異，最終確定為 ResNet101+SGD+Transfer learning 的方案，而訓練出的模型高達 98% 的辨識率。

**關鍵詞：**深度學習、醫學影像、息肉狀脈絡膜血管病變、脈絡膜新生血管、辨識與分類、ResNet101。

## Abstract

The purpose of this paper is to use deep learning to identify and classify polypoidal choroidal vasculopathy (PCV) [1] and choroidal neovascularization (CNV) [2] in medical images. Polypoid choroidal vascular disease, one of the subspecies of age-related macular degeneration, can cause macular hemorrhage and water accumulation in the retina, and is one of the most important causes of vision loss in the elderly. Choroidal neovascularization is also one of the subspecies of age-related macular degeneration, but the mechanism of retinal hemorrhage caused by choroidal neovascularization is different. At present, the two have not yet recognized the unique characteristic genes, so that people can not distinguish between the two by analyzing the genetic correlation, which is not good for early diagnosis and treatment [3]. At present, the common discrimination method is to identify the detailed differences between the two (for example: blood vessel radius, abnormal branch blood vessel network, etc.) through experienced doctors, which is labor-consuming and time-consuming. In this experiment, we used the four architectures of VGG16, ResNet50, ResNet101, ResNet152, plus transfer learning, and used different optimizers and data augmentation experiments to learn the subtle differences between the two images. ResNet101 + SGD + Transfer learning scheme, and the trained model has a recognition rate of up to 98%.

**Keywords :** deep learning, medical images, PCV, CNV, ResNet101.

# 一、介紹

醫學領域的涉及往往需要專業的知識應用，專業的人士進行判斷，受限於高門檻，人力資源缺乏的情形屢見不鮮。而醫學病例的判別，例如醫學影像的部分，更是需要經驗豐富的專業人士才能判斷，而不能透過單純的計算去分析，或是計算的效果不如預期。而其中醫學影像中的一個案例，息肉狀脈絡膜血管病變與脈絡膜新生血管，由於醫學界目前缺乏公認的兩者獨特的基因序列，導致目前的辨識只能透過影像進行分析，更是耗費人力資源。而我們本篇論文目的在於使用深度學習來辨識兩者之間的差異。我們本次實驗中分別採用了 VGG16, ResNet50, ResNet101, ResNet152 四種架構加上遷移式學習並搭配不同優化器與進行資料擴增進行試驗，學習上述兩者影像之間的細微差異，最終確定為 ResNet101+SGD+Transfer learning 的方案，而訓練出的模型高達 98% 的辨識率。

## 二、相關背景知識

**息肉狀脈絡膜血管病變 (polypoidal choroidal vasculopathy, PCV)** 是一種與一般老年性黃斑部病變較不相同的黃斑部病變。它是在脈絡膜層的脈絡膜血管發生不正常的擴大及末端膨大，造成出血、滲漏、積液，進而影響視力。本文將介紹此種黃斑部病變並提出一個病例報告，從診斷到治療討論此疾病與一般老年性黃斑部病變不同的病程及預後。

**脈絡膜新生血管 (choroidal neovascularization CNV)** 這種血管本身是源自於視網膜下的脈絡膜，由脈絡膜長出卻造成視網膜出血，嚴重時甚至玻璃體出血，這種血管所造成的疾病，與一般視網膜本身的疾病是不相同的。這種血管與正常血管最大不同之處，在於血管管壁容易破裂出血，有時甚至會造成水腫及滲出物。這種血管本身若破裂造成出血及滲出物，很容易就會佔據整個黃斑部，也非傳統小光斑雷射所能治療。

## 三、方法

我們的方法主要在於實驗各種不同搭配方案，力求得到最佳辨識率的模型。我們這次實驗主要透過調整以下幾個方向搭配來達到最好的效果。方向分別為以下幾個小節：3.1 dataset 劃分，3.2 參數調整，3.3 模型更換，3.4 遷移式學習，3.5 資料擴增，3.6 驗證方式。

**3.1 dataset 劃分** 我們 dataset 的劃分方式分為兩種：第一種是隨機劃分訓練集與測試集，再將訓練集隨機劃分為訓練集與驗證集，三者比例為，訓練集:驗證集:測試集=8:1:1。第二種是將驗證集與測試集分別以間隔 5 張的頻率從資料集中挑出。因為我們研究後發現，資料集的排序方式會將相近特徵的圖片排序在一起，為了避免隨機抽取導致模型無法訓練到所有特徵，

所以將測試集與驗證集以固定頻率抽取，避免出現隨機集中抽取導致可能的訓練集缺陷。結果（會在實驗中說明）證明，第二種方式的準確率較第一種高。

### 3.2 參數調整 我們的參數調整分成以下幾種：

#### Epochs :

我們的 epoch 次數從 20 → 50 → 100 → 120 → 100 分別調整過，發現在 VGG16 epoch=20 已收斂，而 ResNet 則在 epoch=40 才收斂，但從 test 的結果來看，ResNet 在 epoch=100 時的表現最穩定。

#### Learning rate :

我們的 learning rate 從 0.01 → 0.6 → 0.001 → 0.0001 → 0.001 進行測試，最終結果為 0.001。

#### Optimizer :

我們的 Optimizer 從 SGD → Adam → SGD 進行測試，發現 Adam 較適用於大型訓練集與從較大 learning rate 開始，在小型資料集上效果不好。

#### Loss function :

皆為 Cross Entropy

### 3.3 模型更換 我們的 model 採用以下幾種

VGG16[4]: Visual Geometry Group 16 layers. VGG16 分別為 13 個卷積層與 3 個全連接層，參數量約為 138,360,000。較適用於大型訓練集與單一特徵的訓練上。

ResNet50[5]: Residual neural network. 又名殘差神經網路，指的是在傳統卷積神經網路中加入殘差學習（residual learning）的思想，解決了深層網路中梯度彌散和精度下降（訓練集）的問題，使網路能夠越來越深，既保證了精度，又控制了速度。隨著網路的加深，梯度彌散問題會越來越嚴重，導致網路很難收斂甚至無法收斂。梯度彌散問題目前有很多的解決辦法，包括網路初始標準化，資料標準化以及中間層的標準化（Batch Normalization）等。但是網路加深還會帶來另外一個問題：隨著網路加深，出現訓練集準確率下降的現象。針對這個問題提出了殘差學習的思想：殘差學習的思想，如圖 1，可以把它理解為一個 block，定義如下：

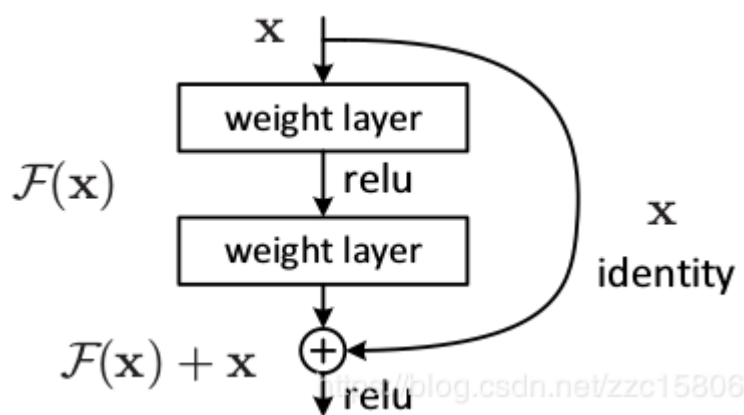


圖 1.  $y = F(x, \{W_i\}) + x$

對於一個神經網路模型，如果該模型是最優的，那麼訓練就很容易將 residual mapping 優化到 0，此時只剩下 identity mapping，那麼無論怎麼增加深度，理論上網路會一直處於最優狀態。因為相當於後面所有增加的網路都會沿著 identity mapping（自身）進行資訊傳輸，可以理解為最優網路後面的層數都是廢掉的（不具備特徵提取的能力），實際上沒起什麼作用。這樣，網路的效能也就不會隨著深度的增加而降低了。常用於影像集的訓練。

ResNet50: 參數量約為 25,560,000，只有 VGG16 的五分之一。

ResNet101: 參數量約為 44,500,000，隱藏層層數達 101 層，比 ResNet50 多了 17 個 convolution block(conv4\_x)。

ResNet152: 參數量約為 60,200,000，隱藏層層數達 152 層。

**3.4 遷移式學習** 遷移式學習簡單來說，就是將原先用於另一項任務的神經網路的模型（稱作 pretrained model）用於另一個新的領域。舉個例子，需要辨別的資料為目標資料 (Target data)，原本既有的資料則稱為來源資料 (Source data)。我們使用的方法叫作 Model fine-tuning，這個使用別人當作初始值，自己再做訓練的過程叫作 fine-tune。但是我們若直接使用 target data fine-tune 出新的模型的話，通常結果會很差。原因是因為 target data 通常數量很少，原來的模型複雜度相對太高，所以訓練出來的結果很容易 over-fitting。因此我們可以將前面幾層 layer 的 pretrain model 保留，代換掉後面幾層的 layer 之後重新訓練模型的參數，便可以拿到一個新的並且符合需求的模型了，而且只訓練後面幾層，參數也較少，較不會發生 over-fitting。

**3.5 資料擴增** 資料擴增適用於小訓練集或是訓練資缺乏的情況下，採用資料擴增的好處之一是可以避免 over fitting。我們採用的資料擴增方式為隨機翻轉與隨機加強對比度，最終結果證明可以確實降低 over fitting，但相對的測試準確率卻大幅下降，所以最終不採用這個調整方案。

**3.6 驗證方式** 驗證我們當模型的方式式採用第一種資料集劃分方式，將訓練集隨機劃分為訓練集與驗證集，並根據驗證集的成功率取得最好的 model，重複 10 次。最終將 10 個 model 分別用於測試集，得到平均準確率 94.38 %

介紹完以上幾種調整方向後，我們會在實驗部分介紹我們搭配了幾種方案來取得最好的效果，先聲明，最終我們最好的實驗方案是: dataset 採用第二種劃分方式 epochs=100, learning rate=0.001, optimizer=SGD, model=ResNet101, Loss function=Cross entropy 加上 transfer learning, 最高準確率為 98.87 %。

## 四、實驗

針對3.2所條列出來的調整方案，我們總共嘗試了6個實驗組並與1個default設定的對照組進行比較。在4.1的部分會提供各實驗的詳細訓練參數及說明，而在4.2中，我們會以表格的方式比較所有實驗的差異並列出準確率。

## 4.1 實驗實作細節

### 4.1.1 Default setting

- 說明：原始的設定是遵照老師所給的範例程式及參數，其中在epoch的部分我們有調整到50，但從結果來看，模型在原始設定的20個epoch便已經收斂。
- 參數設定
  - Model : VGG16
  - Epoch : 20 → 50 → 20
  - Learning rate : 0.01 → 0.6 → 0.01
  - Loss function : cross entropy
  - Optimizer : SGD
- 結果
  - train accuracy: 74.87%
  - test accuracy(cnv): 57.8125%
  - test accuracy(pcv): 76.2500%
  - test accuracy all: 68.0556%

### 4.1.2 Optimizer調整

- 說明：在VGG16的基礎上，我們嘗試使用Adam當作optimizer，但結果卻不如預期，我們推斷可能是因為Adam相較於SGD有著learning rate自我調節機制的功能，通常從較大的learning rate開始訓練起，因此適用於較大的資料集。
- 參數設定
  - Model : VGG16
  - Epoch : 20
  - Learning rate : 0.01
  - Loss function : cross entropy
  - Optimizer : Adam
- 結果
  - train accuracy: 55.23%
  - test accuracy(cnv): 0.0000%
  - test accuracy(pcv): 100.0000%
  - test accuracy all: 55.5556%

### 4.1.3 Model調整

- 說明：會選擇ResNet是因為他本身的結構也是由convolution layer所構成，與VGG的結構差異不大，適用於影像分類，但是殘差網路的設計讓ResNet的訓練效率更佳，故選用ResNet當作我們嘗試的模型，測試結果也的確比VGG16好。
- 參數設定
  - Model : ResNet50
  - Epoch : 50
  - Learning rate : 0.01
  - Loss function : cross entropy
  - Optimizer : SGD



- 結果
- train accuracy: 76.76%
- test accuracy(cnv): 73.4375%
- test accuracy(pcv): 83.7500%
- test accuracy all: 79.1667%

#### 4.1.4 Transfer learning

- 說明：當資料量不足且模型初始權重以隨機的方式給定，則訓練出來的模型權重及參數可能不會是最佳權重，因此借助pretrained weight的方式，讓模型能在一定的基礎上進行訓練。從實驗結果來看，此方提升了模型將近8%以上的準確率。
- 參數設定
  - Model : ResNet50 (Pretrained)+ 1 fully-connected layer
  - Epoch : 50
  - Learning rate : 0.01
  - Loss function : cross entropy
  - Optimizer : SGD
- 結果
  - train accuracy: 90.02%
  - test accuracy(cnv): 87.5000%
  - test accuracy(pcv): 87.5000%
  - test accuracy all: 87.5000%

#### 4.1.5 Data augmentation

- 說明：資料擴增能解決兩大問題，第一個是訓練資料不足的問題，第二個是能避免overfitting的情況發生。而在此部分，我們考量了此次影像的特性，選擇翻轉和調整對比度的方式進行資料擴增。最後的準確率稍微下降了一些，但testing accuracy的表現並不會比train accuracy差，在某種程度上解決了overfitting的問題。但因為整題的成效並不佳，所以最後並未採用data augmentation。
- 參數設定
  - Model : ResNet50 (Pretrained)+ 1 fully-connected layer
  - Data augmentation : 翻轉 + 調整對比
  - Epoch : 50
  - Learning rate : 0.01
  - Loss function : cross entropy
  - Optimizer : SGD
- 結果
  - train accuracy: 72.12%
  - test accuracy(cnv): 84.3750%
  - test accuracy(pcv): 68.7500%
  - test accuracy all: 75.6944%

#### 4.1.6 模型深度加強

- 說明：在ResNet50的基礎上，我們決定嘗試層數更多的ResNet152。為了因應更大型的模型，我們將epoch從50調整到100，並將learning rate從0.01改為0.001，拉長整體的訓練過程，讓模型做更細部的調整。
- 參數設定
  - Model : ResNet152 (Pretrained)+ 1 fully-connected layer
  - Epoch : 100
  - Learning rate : 0.01  $\rightarrow$  0.001
  - Loss function : cross entropy
  - Optimizer : SGD
- 結果
  - train accuracy: 100.00%
  - test accuracy(cnv): 95.3215%
  - test accuracy(pcv): 96.2500%
  - test accuracy all: 95.8333%

#### 4.1.7 模型深度調整(最佳結果)

- 說明：最後我們發現在此資料集下，ResNet101能有最佳的表現。其中我們也嘗試將epoch上調，但事實上反而讓準確率下降。在learning rate的部分，我們也發現下調到0.0001時，容易讓模型落入了局部最佳解，故最後還是設定為0.001。
- 參數設定
  - Model : ResNet101 (Pretrained)+ 1 fully-connected layer
  - Epoch : 100  $\rightarrow$  120  $\rightarrow$  100
  - Learning rate : 0.001  $\rightarrow$  0.0001  $\rightarrow$  0.001
  - Loss function : cross entropy
  - Optimizer : SGD
- 結果
  - train accuracy: 100.00%
  - test accuracy(cnv): 98.4375%
  - test accuracy(pcv): 96.7500%
  - test accuracy all: 98.6111%

```
Accuracy of   cnv : 98.4375 %  
Accuracy of   pcv : 98.7500 %  
  
Accuracy on the ALL test images: 98.6111 %
```

## 4.2 比較總表

Model	pretrained	optimizer	epoch	Learning rate	Training accuracy	Testing accuracy	備註
VGG16 (default)	No	SGD	20	0.01	74.87%	68.0556%	
	No	Adam	20	0.01	55.23%	55.556%	
ResNet50	No	SGD	50	0.01	76.76%	79.1667%	
	Yes	SGD	50	0.01	90.02%	87.5000%	
	Yes	SGD	50	0.01	72.12%	75.6944%	資料擴增
ResNet152	Yes	SGD	50	0.01	97.76%	91.6667%	
	Yes	SGD	100	0.001	100.0%	95.8333%	
ResNet101	Yes	SGD	100	0.001	100.0%	98.6111%	最佳

## 五、結論

本次的醫學影像專題中主要是針對視網膜中兩種病理現象進行辨識，由於訓練資料不算充足，所以適合使用參數量較小的模型進行訓練，其中 ResNet 的 residual block 使得它在這次的專題中能有相當優異的表現。而在資料分布的部分，這次的資料有相同病人的圖片聚集在一起的現象，導致訓練和測試資料的切分方式也會影響這次的實驗結果。

總體而言，我們這次的貢獻大致有下列幾項：

1. 觀察圖片在資料夾中的分布，並提供不同劃分 dataset 的方式，最後加以驗證。
2. 分析不同模型的優劣，選擇出適當的模型，並將準確率提升到最高 98.6111%
3. 針對不同的參數進行調整，並找出穩定且表現優異的參數組合。

## 六、參考

- [1] Antonio P. Ciardella, MD, Irene M. Donsoff, MD, Sheau J. Huang, MD, Danielle L. Costa, MD, and Lawrence A. Yannuzzi, MD, “Polypoidal Choroidal Vasculopathy”, SURVEY OF OPHTHALMOLOGY, January–February 2004, Pages 25-37
- [2] HANS E. GROSSNIKLAUS, MD, AND W. RICHARD GREEN, MD, “Choroidal neovascularization”, American Journal of Ophthalmology, March 2004, Pages 496-503
- [3] Belinda K. Cornes; Wan Ting Tay; Xueling Sim; Peter D. Cackett; Eranga N. Vithana; E. Shyong Tai; Tin Aung; Tien Y. Wong, “GWAS of Polypoidal Choroidal Vasculopathy (PCV) and Choroidal Neovascularization (CNV) in a Chinese Cohort from Singapore”, ARVO Annual Meeting, April 2011
- [4] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, Alan L. Yuille, “Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs”, arXiv, 2014
- [5] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, Alexander A. Alemi, “Inception-v4,

Inception-ResNet and the Impact of Residual Connections on Learning”, Google, AAAI, 2015