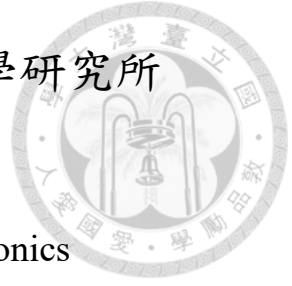


國立臺灣大學電機資訊學院光電工程學研究所

碩士論文

Graduate Institute of Photonics and Optoelectronics
College of Electrical Engineering and Computer Science
National Taiwan University
Master's Thesis



應用自監督學習及半監督聚類於視網膜光學同調斷層
掃描影像探討潛在之亞健康樣本

Exploring potential sub-healthy eyes in retina OCT images
using self-supervised learning and semi-supervised
clustering

張翔滄

Hsiang-Yu Chang

指導教授：黃升龍 博士

Advisor: Sheng-Lung Huang, Ph.D.

中華民國 114 年 7 月

July 2025

致謝

首先我要感謝指導教授黃升龍老師，在每週的討論與報告中，老師不僅引導我釐清研究方向，也大幅提升了我的表達與報告能力。每當我在研究過程中遇到困惑時，老師總提供寶貴的建議讓我獲得啟發，並且在每次與醫師報告前老師總會提醒我報告的重點與技巧。在論文撰寫期間，老師更是不辭辛勞地多次協助我修改初稿，讓本論文得以逐步完善，深感感激。還需要特別感謝謝易庭醫師，即使我對於視網膜影像並不了解，但每當我遇到問題向謝醫師請教時，謝醫師也總是耐心且詳細的回答我，並且還抽空與我討論並分享對於我挑選出之潛在亞健康樣本之看法。

接下來我要感謝我的父母，非常感謝他們對於我讀碩士的支持，讓我能夠專心的完成研究，不需要額外擔心其他事務。

我還要感謝李劭允學長，非常感謝學長在畢業後仍與我交接各項學長所使用的程式並跟我介紹各個資料集的內容，讓我在一開始時能較快的上手。我還要感謝李宇昂學長，在我寫程式的過程中遇到困難，學長也都很願意跟我分享他的見解，以及教導我該如何修改並了解程式。我也要感謝以佳學姊及宥亘學長，感謝兩位學長姊跟我分享對於進度的安排及建議。還要感謝兆昇、以樂、子昕、威丞，在修課時相互學習，並督促對方每周要報告。

最後我要感謝瀚陞、阿煒、尻哥、其易、蘇菲亞，感謝督促我做實驗做報告，並且在我煩悶時會陪我聊天、出門、攀岩，讓我又能重新找到動力繼續完成研究，甚至有時即使對於我的研究內容並不了解，但還是願意陪我練習報告，讓我能與老師的討論中表達得更加清晰。

中文摘要



光學同調斷層掃描(Optical Coherence Tomography; OCT)是一種高解析度的非侵入式成像技術，具備快速成像與優異解析力，廣泛應用於眼科臨床診斷。其產生之高品質影像不僅有助於醫師進行病灶判讀，也能提升卷積神經網路(Convolutional Neural Network; CNN)對影像特徵的學習效率，進而優化模型表現。根據先前研究，CNN 已成功應用於視網膜 OCT 影像之疾病分類任務，證實其於健康與病理影像判別上的可行性。

為進一步提升特徵表徵能力，本研究採用自監督學習(Self-Supervised Learning; SSL)中的 SimCLR 演算法進行模型預訓練。與傳統監督式學習不同，SimCLR 無需人工標註，能利用大規模未標註資料進行學習，降低標註成本，並可使用與下游任務同源之醫學影像，較不受 ImageNet 自然影像資料集之領域差異影響。本研究使用 19464 張 OCT 影像進行 SimCLR 預訓練，相較之下，ImageNet 資料集則包含 1280 萬張自然影像。

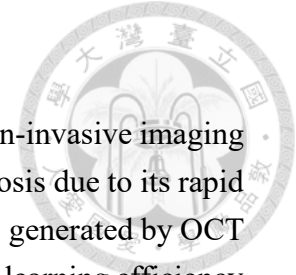
實驗結果顯示，雖然以 SimCLR 預訓練之模型在多項分類任務中皆能順利完成目標，整體表現卻常略遜於以 ImageNet 資料集預訓練之模型，推測與預訓練資料集規模上的巨大差異有關。此外，訓練初期驗證準確率高於訓練集，收斂後則反轉。此現象是因為訓練資料為提升泛化能力會隨機水平翻轉，而驗證資料則未進行增強；隨著訓練次數增加，模型逐漸熟悉訓練資料並有效擬合，因此訓練集表現提升，最終準確率超越驗證集。然而在健康與亞健康分類任務中，無論採用哪種預訓練策略，模型平均準確率皆約為 66%，推測部分亞健康樣本因對側眼還未確診為黃斑前膜(Epiretinal Membrane; ERM)而誤標為健康，進而影響模型學習。

為進一步探討資料集中潛在亞健康樣本之分布，本研究結合無監督聚類分析 K-means 與半監督聚類分析 Semi-supervised Deep Embedded Clustering(SDEC)進行分析。於 SDEC 中僅提供亞健康樣本之標註，引導模型學習其潛在特徵。雖兩種聚類方法皆能在視覺化中呈現清晰分群，但實際標註影像卻呈現沒有規則之分布。然而，經多次實驗統計結果可知，SDEC 模型在重複挑選潛在亞健康樣本上展現較高一一致性。經由與醫師之討論確認，部分被挑選之樣本中央凹弧度較平緩，與先前研究所述亞健康特徵相符。

綜上所述，本研究期望透過 SDEC 分群方式，有效輔助找出健康標註資料中之潛在亞健康樣本，進而提升 CNN 模型於健康與亞健康分類任務中的效能。未來若能搭配臨床追蹤資料驗證，將有助於早期辨識黃斑前膜病變風險，達到預防與精準診療之目標。

關鍵字：光學同調斷層掃描、卷積神經網路、自監督學習、半監督聚類分析、黃斑前膜亞健康

Abstract



Optical Coherence Tomography (OCT) is a high-resolution, non-invasive imaging technique that has been widely applied in ophthalmic clinical diagnosis due to its rapid image acquisition and excellent resolution. The high-quality images generated by OCT not only assist physicians in identifying lesions but also enhance the learning efficiency of Convolutional Neural Networks (CNN) in extracting image features, thereby improving model performance. According to previous studies, CNNs have been successfully applied to retinal OCT image classification tasks, demonstrating their feasibility in distinguishing between healthy and pathological cases.

To further enhance feature representation, this study adopts SimCLR, a Self-Supervised Learning (SSL) algorithm, for model pretraining. Unlike traditional supervised learning, SimCLR does not require manual annotation and can learn from large-scale unlabelled datasets, reducing annotation costs. Furthermore, it enables the use of medical images that are consistent with downstream tasks, minimizing domain gaps compared to pretraining on natural image datasets like ImageNet. In this study, 19464 OCT images were used for SimCLR pretraining, whereas the ImageNet dataset contains approximately 12.8 million natural images.

Experimental results show that although the SimCLR-pretrained model successfully accomplishes multiple classification tasks, its overall performance is generally slightly inferior to that of the ImageNet-pretrained model. This discrepancy is presumed to stem from the significant difference in dataset scale between the two pretraining strategies. In addition, validation accuracy was higher than training accuracy during the early training phase but reversed as the model converged. This was due to the use of random horizontal flipping in the training data to enhance generalization, while no augmentation was applied to the validation data. As training progressed, the model gradually became more familiar with the training data and fit it more effectively, resulting in higher training accuracy in the later stages. However, in the classification task of distinguishing between healthy and sub-healthy cases, both pretraining methods yield similar average accuracies of approximately 66%. This may be due to some sub-healthy samples being mislabeled as healthy, as their contralateral eye had not yet been clinically diagnosed with Epiretinal Membrane (ERM), thereby affecting model learning.

To further investigate the distribution of potential sub-healthy samples in the dataset, this study employs both unsupervised clustering (K-means) and semi-supervised clustering (Semi-supervised Deep Embedded Clustering, SDEC). In the SDEC framework, only labels of sub-healthy samples are provided to guide the model in learning latent features. Although both clustering methods produce visually distinct

groupings, the labeled samples appear randomly distributed. Nevertheless, statistical results from repeated experiments indicate that the SDEC model demonstrates higher consistency in repeatedly identifying potential sub-healthy samples. Discussions with ophthalmologists confirmed that some of the selected samples exhibited a flatter foveal contour, which aligns with previously reported characteristics of sub-healthy cases.

In summary, this study aims to utilize the SDEC clustering method to effectively assist in identifying potential sub-healthy samples within the healthy-labeled dataset, thereby improving CNN performance in healthy versus sub-healthy classification tasks. If validated by future longitudinal clinical data, this approach may contribute to the early detection of ERM risk and support the goals of preventive and precision medicine.

Keywords: Optical Coherence Tomography, Convolutional Neural Network, Self-Supervised Learning, Semi-Supervised Clustering, Sub-healthy Samples of Epiretinal Membrane

目次



致謝.....	I
中文摘要.....	II
Abstract.....	III
目次.....	V
圖次.....	VII
表次.....	IX
Chapter 1 緒論.....	1
Chapter 2 頻域式光學同調斷層掃描及視網膜介紹.....	3
2.1 光學同調斷層掃描.....	3
2.1.1 光學同調斷層掃描原理.....	3
2.1.2 頻域式光學同調斷層掃描.....	4
2.1.3 海德堡光學同調斷層掃描顯微儀.....	6
2.2 視網膜介紹.....	6
2.2.1 視網膜之結構.....	6
2.2.2 黃斑部與中央凹之功能.....	9
2.3 視網膜疾病.....	10
2.3.1 黃斑前膜.....	10
2.3.2 黃斑前膜之亞健康定義.....	10
Chapter 3 研究方法及資料集.....	12
3.1 卷積神經網絡基本架構及殘差網絡架構介紹.....	12
3.2 自監督學習.....	15
3.2.1 自監督學習與監督學習之介紹與比較.....	15
3.2.2 SimCLR 流程及應用.....	16
3.3 聚類方式.....	17
3.3.1 無監督聚類.....	18
3.3.2 半監督聚類.....	18
3.4 資料集與前處理.....	20
3.4.1 資料集之來源及數量.....	20
3.4.2 影像之前處理.....	20
Chapter 4 健康、亞健康與不健康眼睛之分類.....	22
4.1 實驗設計.....	22
4.1.1 模型架構及預訓練方式之介紹.....	22
4.1.2 訓練設定及評估指標.....	24
4.2 分類任務一：健康與不健康之分類.....	25

4.3	分類任務二：黃斑前膜與其他疾病之分類.....	30
4.4	分類任務三：健康與黃斑前膜之分類.....	35
4.5	分類任務四：健康與亞健康之分類.....	40
4.6	各項分類任務之分析與討論.....	44
Chapter 5	聚類分析與潛在亞健康樣本之探討.....	46
5.1	聚類分析設計與方法.....	46
5.2	無監督聚類分析.....	47
5.3	半監督聚類分析.....	52
5.4	聚類結果之分析與討論.....	56
Chapter 6	結果與未來展望.....	61
6.1	結論.....	61
6.2	未來展望.....	62
Reference	64
附錄 1	三種不同預訓練模型在 K-means 分群中所選之潛在亞健康樣本.....	67
	ImageNet Pre-train Model.....	67
	SimCLR Pre-train Model.....	68
	OCT Pre-train Model.....	70
附錄 2	三種不同預訓練模型在 SDEC 分群中所選之潛在亞健康樣本.....	72
	ImageNet Pre-train Model.....	73
	SimCLR Pre-train Model.....	75
	OCT Pre-train Model.....	77

圖次

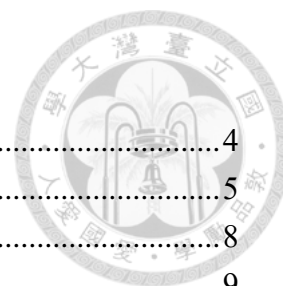


圖 2.1 麥克森干涉儀基本架構.....	4
圖 2.2 SD-OCT 架構示意圖.....	5
圖 2.3 視網膜十層構造圖.....	8
圖 2.4 感光細胞示意圖.....	9
圖 2.5 黃斑部之 OCT 影像.....	9
圖 2.6 健康樣本示意圖.....	10
圖 2.7 亞健康樣本示意圖.....	11
圖 3.1 卷積層示意圖.....	13
圖 3.2 ReLU function.....	13
圖 3.3 卷積神經網絡基本架構圖.....	14
圖 3.4 殘差學習示意圖.....	14
圖 3.5 監督學習與自監督學習示意圖.....	16
圖 3.6 SimCLR 訓練流程.....	17
圖 3.7 SDEC 訓練流程.....	19
圖 3.8 視網膜 OCT 原始影像.....	21
圖 3.9 視網膜 OCT 裁切後之影像.....	21
圖 4.1 在 SimCLR 自監督學習預訓練中訓練集與驗證集對影像增強產生偽標籤 對分類之準確率.....	23
圖 4.2 在 SimCLR 自監督學習預訓練中訓練集與驗證集對影像增強產生偽標籤 對分類之 Loss.....	24
圖 4.3 Cross Validation 示意圖.....	25
圖 4.4 ImageNet Pre-train Model 在健康眼睛與不健康眼睛(包含 ERM)分類任 務中訓練集與驗證集準確率與 Epoch 做圖結果.....	26
圖 4.5 ImageNet Pre-train Model 在健康眼睛與不健康眼睛(包含 ERM)分類任 務中訓練集與驗證集 Loss 與 Epoch 做圖結果.....	27
圖 4.6 SimCLR Pre-train Model 在健康眼睛與不健康眼睛(包含 ERM)分類任 務中訓練集與驗證集準確率與 Epoch 做圖結果.....	28
圖 4.7 SimCLR Pre-train Model 在健康眼睛與不健康眼睛(包含 ERM)分類任 務中訓練集與驗證集 Loss 與 Epoch 做圖結果.....	28
圖 4.8 ImageNet Pre-train Model 在黃斑前膜與其他疾病分類任務中訓練集與驗 證集準確率與 Epoch 做圖結果.....	32
圖 4.9 ImageNet Pre-train Model 在黃斑前膜與其他疾病分類任務中訓練集與驗 證集 Loss 與 Epoch 做圖結果.....	32
圖 4.10 SimCLR Pre-train Model 在黃斑前膜與其他疾病分類任務中訓練集與驗 證集準確率與 Epoch 做圖結果.....	33

圖 4.11 SimCLR Pre-train Model 在黃斑前膜與其他疾病分類任務中訓練集與驗證集 Loss 與 Epoch 做圖結果.....	33
圖 4.12 ImageNet Pre-train Model 在健康與黃斑前膜分類任務中訓練集與驗證集準確率與 Epoch 做圖結果	37
圖 4.13 ImageNet Pre-train Model 在健康與黃斑前膜分類任務中訓練集與驗證集 Loss 與 Epoch 做圖結果.....	37
圖 4.14 SimCLR Pre-train Model 在健康與黃斑前膜分類任務中訓練集與驗證集準確率與 Epoch 做圖結果	38
圖 4.15 SimCLR Pre-train Model 在健康與黃斑前膜分類任務中訓練集與驗證集 Loss 與 Epoch 做圖結果.....	38
圖 4.16 ImageNet Pre-train Model 在健康與亞健康分類任務中訓練集與驗證集準確率與 Epoch 做圖結果	42
圖 4.17 ImageNet Pre-train Model 在健康與亞健康分類任務中訓練集與驗證集 Loss 與 Epoch 做圖結果.....	42
圖 4.18 SimCLR Pre-train Model 在健康與亞健康分類任務中訓練集與驗證集準確率與 Epoch 做圖結果	43
圖 4.19 SimCLR Pre-train Model 在健康與亞健康分類任務中訓練集與驗證集 Loss 與 Epoch 做圖結果.....	43
圖 4.20 OCT Pre-train Model 在健康與亞健康分類任務中訓練集與驗證集準確率與 Epoch 做圖結果	44
圖 4.21 OCT Pre-train Model 在健康與亞健康分類任務中訓練集與驗證集 Loss 與 Epoch 做圖結果	44
圖 5.1 混淆矩陣示意圖.....	47
圖 5.2 ImageNet Pre-train Model 無監督聚類分析預測分群與實際分群結果	48
圖 5.3 SimCLR Pre-train Model 無監督聚類分析預測分群與實際分群結果.....	48
圖 5.4 OCT Pre-train Model 無監督聚類分析預測分群與實際分群結果	48
圖 5.5 ImageNet Pre-train Model 半監督聚類分析預測分群與實際分群結果	53
圖 5.6 SimCLR Pre-train Model 半監督聚類分析預測分群與實際分群結果.....	53
圖 5.7 OCT Pre-train Model 半監督聚類分析預測分群與實際分群結果	53
圖 5.8 潛在亞健康樣本之影像.....	58
圖 5.9 ImageNet Pre-train Model 經健康與亞健康分類訓練後對潛在亞健康樣本之 Grad-CAM 影像(0_00092684_11.png).....	58
圖 5.10 SimCLR Pre-train Model 經健康與亞健康分類訓練後對潛在亞健康樣本之 Grad-CAM 影像(0_00092684_11.png).....	58
圖 5.11 OCT Pre-train Model 經健康與亞健康分類訓練後對潛在亞健康樣本之 Grad-CAM 影像(0_00092684_11.png).....	59
圖 5.12 SDEC 訓練中之配對限制指示矩陣	60

表次



表 4.1 本實驗中所使用之模型架構.....	22
表 4.2 健康與不健康之分類任務中各模型之超參數設定.....	25
表 4.3 ImageNet Pre-train Model 在健康眼睛與不健康眼睛(包含 ERM)分類任務中測試集之表現結果.....	26
表 4.4 SimCLR Pre-train Model 在健康眼睛與不健康眼睛(包含 ERM)之分類任務中測試集之表現結果.....	27
表 4.5 ImageNet Pre-train Model 及 SimCLR Pre-train Model 經健康與不健康分類任務訓練後最後一層對健康影像之 Grad-CAM 影像.....	29
表 4.6 ImageNet Pre-train Model 及 SimCLR Pre-train Model 經健康與不健康分類任務訓練後最後一層對不健康影像之 Grad-CAM 影像.....	30
表 4.7 黃斑前膜與其他疾病之分類任務中各模型之超參數設定.....	31
表 4.8 ImageNet Pre-train Model 在黃斑前膜與其他疾病分類任務中測試集之表現結果.....	31
表 4.9 SimCLR Pre-train Model 在黃斑前膜與其他疾病分類任務中測試集之表現結果.....	33
表 4.10 ImageNet Pre-train Model 及 SimCLR Pre-train Model 經黃斑前膜與其他疾病分類任務訓練後最後一層對黃斑前膜影像之 Grad-CAM 影像.....	34
表 4.11 ImageNet Pre-train Model 及 SimCLR Pre-train Model 經黃斑前膜與其他疾病分類任務訓練後最後一層對其他疾病影像之 Grad-CAM 影像.....	35
表 4.12 健康與黃斑前膜之分類任務中各模型之超參數設定.....	36
表 4.13 ImageNet Pre-train Model 在健康與黃斑前膜分類任務中測試集之表現結果.....	36
表 4.14 SimCLR Pre-train Model 在健康與黃斑前膜分類任務中測試集之表現結果.....	38
表 4.15 ImageNet Pre-train Model 及 SimCLR Pre-train Model 經健康與黃斑前膜分類任務訓練後最後一層對健康影像之 Grad-CAM 影像.....	39
表 4.16 ImageNet Pre-train Model 及 SimCLR Pre-train Model 經健康與黃斑前膜分類任務訓練後最後一層對黃斑前膜影像之 Grad-CAM 影像.....	40
表 4.17 健康與黃斑前膜之分類任務中各模型之超參數設定.....	41
表 4.18 ImageNet Pre-train Model 在健康與亞健康分類任務中測試集之表現結果.....	41
表 4.19 SimCLR Pre-train Model 在健康與亞健康分類任務中測試集之表現結果.....	42
表 4.20 OCT Pre-train Model 在健康與亞健康分類任務中測試集之表現結果....	43
表 5.1 ImageNet Pre-train Model 在 Healthy vs Sub-healthy(Female)資料集無監督	

聚類分析結果.....	49
表 5.2 ImageNet Pre-train Model 在無監督聚類分析選出潛在亞健康樣本統計結果.....	49
表 5.3 SimCLR Pre-train Model 在 Healthy vs Sub-healthy(Female)資料集無監督聚類分析結果.....	50
表 5.4 SimCLR Pre-train Model 在無監督聚類分析選出潛在亞健康樣本統計結果.....	50
表 5.5 OCT Pre-train Model 在 Healthy vs Sub-healthy(Female)資料集無監督聚類分析結果.....	51
表 5.6 OCT Pre-train Model 在無監督聚類分析選出潛在亞健康樣本統計結果..	51
表 5.7 半監督聚類分析各模型之超參數設定.....	52
表 5.8 ImageNet Pre-train Model 在 Healthy vs Sub-healthy(Female)資料集半監督聚類分析結果.....	54
表 5.9 ImageNet Pre-train Model 在半監督聚類分析選出潛在亞健康樣本統計結果.....	54
表 5.10 SimCLR Pre-train Model 在 Healthy vs Sub-healthy(Female)資料集半監督聚類分析結果.....	55
表 5.11 SimCLR Pre-train Model 在半監督聚類分析選出潛在亞健康樣本統計結果.....	55
表 5.12 OCT Pre-train Model 在 Healthy vs Sub-healthy(Female)資料集半監督聚類分析結果.....	55
表 5.13 OCT Pre-train Model 在半監督聚類分析選出潛在亞健康樣本統計結果	56



Chapter 1 緒論

光學同調斷層掃描(Optical Coherence Tomography; OCT)是一種非侵入式且具有高解析度的成像技術，透過入射光與反射光之間的干涉現象，可構建出生物組織的橫截面影像。OCT 具備成像速度快、高解析度等優勢，因此在眼科領域中獲得廣泛應用，尤其適用於視網膜結構的觀察與黃斑部病變的診斷。由於 OCT 所產生的影像相比彩色眼底圖擁有更高的解析度及層次細節，不僅可協助臨床醫師進行更精確的診斷與治療評估，也為深度學習技術提供了極佳的資料來源，使模型能夠學習更具語意意涵的視覺特徵。本研究將於第二章中介紹 OCT 技術的成像原理，並進一步探討視網膜之解剖結構與功能。

近年來，深度學習(Deep Learning)在各個領域皆展現出突破性的成果，於醫學領域中亦逐漸成為輔助診斷與決策的重要工具。其中，卷積神經網絡(Convolutional Neural Network; CNN)因其在影像處理任務中的卓越表現，被廣泛應用於各類醫學影像分析任務。透過輸入具標註(Label)之影像資料進行監督式訓練，CNN 得以自動學習影像中的顯著特徵，進而獲得影像與診斷結果之間的對應關係，輔助臨床醫師做出更為準確的診斷決策。此外，隨著模型可視化技術(如 Grad-CAM)的進步，研究者能進一步觀察模型在做出預測時所關注之影像區域，使深度學習模型更具可解釋性與臨床信任度。

然而，深度學習模型之效能高度依賴於資料集之品質與規模。在醫學影像分析中，雖然本研究透過與台大醫院眼科部謝易庭醫師合作，取得了大量視網膜 OCT 影像，但若要使模型能針對特定疾病或病變進行準確分類，仍需針對每張影像進行詳盡且具一致性的人工標註，此步驟不僅極度耗時，亦需仰賴具專業背景之臨床醫師，實務上成本極高。因此，如何在減少標註負擔的同時仍能有效發展具泛化能力之模型，成為當前醫學影像領域的重要挑戰之一。

為了解決上述問題，本研究導入自監督學習(Self-supervised Learning; SSL)進行模型預訓練。自監督學習是一種不依賴人工標註資訊的學習策略，能夠自己產生偽標籤(Pseudo Label)，讓模型從原始影像資料中自我學習影像的結構特性與語意分布。SimCLR 為一種具代表性的自監督對比學習方法，透過資料增強(Data Augmentation)與對比損失函數(Contrastive Loss)，使模型學習將不同影像樣本拉開距離、將同一樣本的不同增強版本拉近，進而學習到對影像具區分性的表徵。相較於傳統以自然影像(如 ImageNet)為基礎的遷移學習，自監督學習能直接使用醫學領域中同源的未標註資料，降低領域不一致問題。本研究將於第三章中詳細說明不同模型之預訓練方式、SimCLR 之訓練流程、使用資料，並在第四章中呈現其於分類任務上的表現與分析。

除分類模型訓練外，本研究亦進一步針對潛在亞健康樣本進行探討。在本研究中，亞健康樣本被定義為：對側眼已臨床確診為黃斑前膜(Epiretinal Membrane;

ERM), 但該眼尚未出現病變之視網膜 OCT 影像。然而在實際人工標註流程中, 若該眼之對側眼尚未發現病變, 仍可能被標為「健康」, 進而使資料集中混入潛在具有病理風險之亞健康影像, 對後續模型學習產生干擾。為補足人工標註之不足與識別盲點, 本研究採用兩種聚類策略進行分析。其一為無監督聚類分析中的 K-means 演算法, 該方法可根據樣本在特徵空間中的相似性自動分群; 其二為半監督式聚類分析(Semi-supervised Deep Embedded Clustering; SDEC), 僅提供亞健康樣本的標註資訊, 引導模型學習其特徵表徵並完成分群任務。透過這些方法, 本研究期望從標註為健康之資料中, 辨識出潛藏的亞健康樣本, 進一步補足臨床資料之盲點。本論文將於第五章呈現聚類結果與潛在亞健康樣本之探討, 並於第六章中整合分類與聚類任務之實驗分析以及外來展望。



Chapter 2 頻域式光學同調斷層掃描及視網膜介紹

在此章中將會對光學同調斷層掃描技術之原理、視網膜的結構、相關疾病以及本實驗中對於黃斑前膜亞健康樣本的定義作介紹。

2.1 光學同調斷層掃描

在此節中會說明頻域式及時域式光學同調斷層掃描之原理並對本實驗收集資料所使用的海德堡光學同調斷層掃描顯微儀系統做介紹。

2.1.1 光學同調斷層掃描原理

光學同調斷層掃描(Optical Coherence Tomography; OCT)技術是利用光的同調性來成像，而其基本架構使用麥克森干涉儀(Michelson Interferometer)如圖 2.1 所示，光束在經過一個 50/50 分光鏡後，分為兩道光並分別打到樣本及參考端反射鏡後重新回到 50/50 分光鏡後，彼此干涉重新組合成一道光，打到偵測器後即可觀察干涉條紋。其中 OCT 可以分為時域式 OCT(Time Domain OCT; TD-OCT)跟頻域式 OCT(Fourier Domain OCT; FD-OCT)。

FD-OCT 與 TD-OCT 最大的不同之處在於 FD-OCT 不需要移動參考端來獲得樣品不同深度的影像資訊，而是在偵測器收到干涉信號後對其做傅立葉轉換以此來獲得樣品在不同深度的資訊。因此 FD-OCT 相比 TD-OCT 更為適合用在掃描眼睛上[1]。

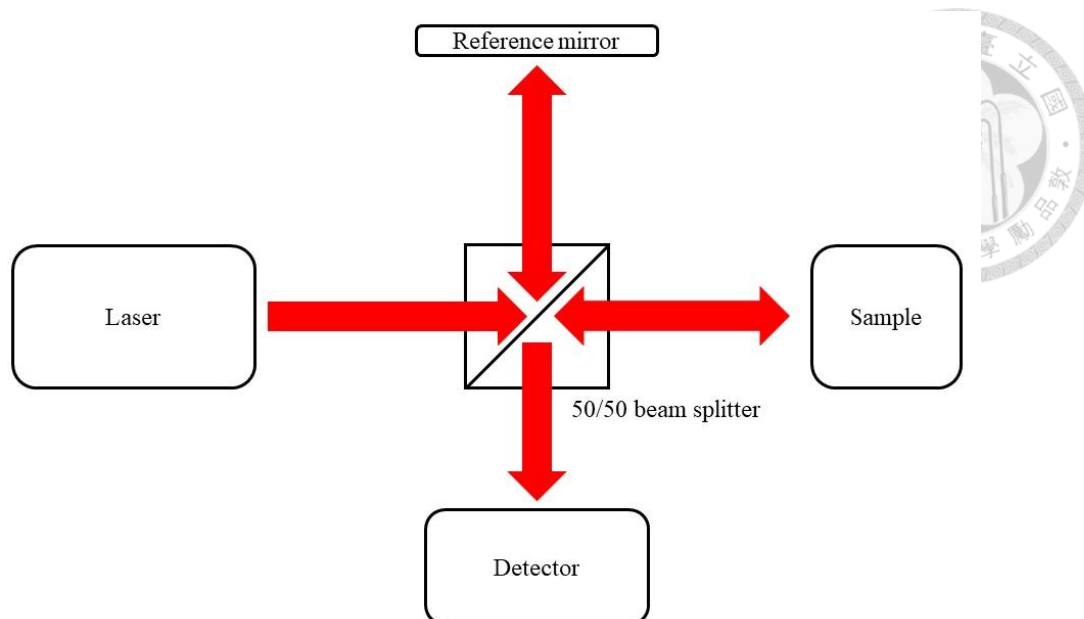


圖 2.1 麥克森干涉儀基本架構

2.1.2 頻域式光學同調斷層掃描

FD-OCT 可以細分為空間域式 OCT(Spectral Domain OCT; SD-OCT)與掃頻式 OCT(Swept Source OCT; SS-OCT)[2]。SD-OCT 通常使用寬頻光源與光譜儀來獲得整個干涉光譜，藉由傅立葉轉換重建深度方向的影像；而 SS-OCT 則使用波長掃描雷射(Swept Source Laser)依序掃描不同波長，並以單一光電偵測器(Single Photodetector)接收訊號[3]。

SD-OCT 獲得樣本不同深度的資訊並不會移動參考端反射鏡而是雷射打入樣本後因為在不同深度有不一樣的折射率而產生不同的反射光(2.1)，

$$R = \left[\frac{n_1 - n_2}{n_1 + n_2} \right]^2 \quad (2.1)$$

R 為反射率； n_1 、 n_2 分別代表樣本在不同深度下所對應的折射率而這些從樣本端及參考端打回來的反射光回到 50/50 分光鏡後彼此干涉，再由光譜儀獲得整個干涉光譜，最終對干涉光譜作傅立葉轉換後得到樣本在不同深度的資訊。

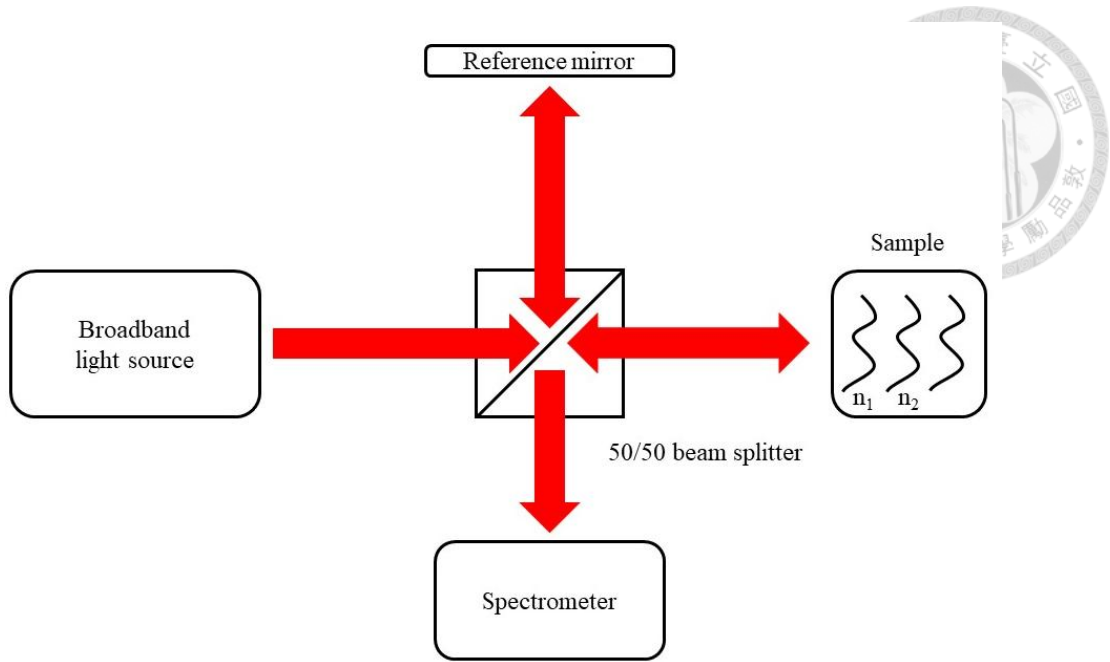


圖 2.2 SD-OCT 架構示意圖

如圖 2.2 所示，一寬頻光在經過一個 50/50 分光鏡後，分為兩道光並分別打到樣本及參考端反射鏡後重新回到 50/50 分光鏡後彼此干涉重新組合成一道光，再由光譜儀接收。假設參考端及樣本端反射光電場分別為 E_R 、 E_S 如式(2.2)、(2.3) 所示。

$$E_R = G(k)a_R e^{i2k\Delta r} \quad (2.2)$$

$$E_S = G(k) \int_{z_0}^{\infty} a(z) e^{i2kn(z)(\Delta r+z)} dz \quad (2.3)$$

$G(k)$ 為寬頻光光譜； a_R 為參考端反射光電場強度； $a(z)$ 為樣本端在不同深度之反射光電場強度； Δr 為參考端與 50/50 分光鏡之移動距離； z 為參考端在初始位置 ($\Delta r = 0$) 時參考臂與樣本臂之間的光程差

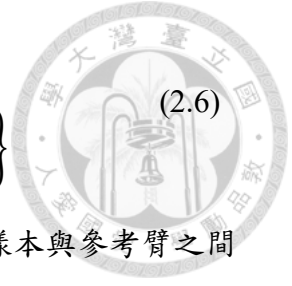
$$I(k) = G(k)^2 \left\{ a_R e^{i2k\Delta r} + \int_{z_0}^{\infty} a(z) e^{i2kn(z)(\Delta r+z)} dz \right\}^2 \quad (2.4)$$

$I(k)$ 為光譜儀所接收道的干涉信號強度，將(2.4)展開後可以得到(2.5)

$$I(k) = G(k)^2 \left\{ (a_R e^{i2k\Delta r})^2 + 2 \int_{z_0}^{\infty} a(z) \cos[2kn(z)(\Delta r+z)] dz + \int_{z_0}^{\infty} a(z) e^{i2kn(z)(\Delta r+z')} dz' \int_{z_0}^{\infty} a(z) e^{i2kn(z)(\Delta r+z)} dz \right\} \quad (2.5)$$

其中又因為 SD-OCT 並不會移動參考端反射鏡，因此 $\Delta r = 0$ ，將 $\Delta r = 0$ 帶回(2.5) 後可以得到(2.6)

$$I(k) = G(k)^2 \left\{ a_R^2 + 2 \int_{z_0}^{\infty} a(z) \cos[2kn(z)z] dz + \int \int_{z_0}^{\infty} a(z) a(z') e^{i2kn(z)(z-z')} dz dz' \right\} \quad (2.6)$$



在(2.6)中 a_R^2 為直流項(DC Term)； $2 \int_{z_0}^{\infty} a(z) \cos[2kn(z)z] dz$ 為樣本與參考臂之間的干涉項(Cross-correlation)； $\int \int_{z_0}^{\infty} a(z) a(z') e^{i2kn(z)(z-z')} dz dz'$ 為樣本中不同深度之間的自干涉項(Auto-correlation)。

由此可知，樣本中不同位置因具有不同的相位差，將產生對應頻率的干涉訊號。除了樣本內部各介面所引起的干涉訊號外，樣本與參考臂之間不同的光程差也會導致相位變化。當光程差較大時，所對應的干涉訊號頻率亦較高；反之，光程差較小則產生較低頻之干涉訊號。因此，透過傅立葉轉換將頻域干涉訊號轉換至時域(2.7)，即可獲得樣本在不同深度的資訊。

$$I(t) = \text{FFT} \{I(k)\} \quad (2.7)$$

2.1.3 海德堡光學同調斷層掃描顯微儀

在本實驗中皆是使用海德堡公司的譜域式光學同調斷層掃描儀(SD-OCT)機型為 Heidelberg Spectralis OCT，光源為波長 870 nm 之超發光二極體(Super Luminescence Diode; SLD)、輸出為 1.2 mW 之連續波；橫向及軸向解析度分別為 14 μm 及 7 μm ；軸向掃描深度約為 1.9 mm，影像包含 469 個像素點。此系統包含兩種掃描模式：

- (1) 高解析度模式：橫向解析度達 6 μm ，軸向解析度為 3.9 μm 。
- (2) 高速模式：橫向解析度為 11 μm ，軸向解析度亦為 3.9 μm 。

在本實驗中所使用的影像皆為使用高速模式獲得。

2.2 視網膜介紹

視網膜位在眼球的最內層於脈絡膜(Choroid)與玻璃體(Vitreous Body)之間，主要由光感受器細胞和膠質細胞組成。視網膜最主要的功能即為將接收到的光信號轉成電信號使人類可以看到影像[4]。

2.2.1 視網膜之結構

視網膜的結構自玻璃體面至脈絡膜方向共可分為十層如圖 2.3 所示，各層結構依序如下：[4—6]：

1. 內界膜(Inner Limiting Membrane):

此層位於玻璃體與視網膜之間的界面，用以維持內層視網膜結構的完整性。

2. 神經纖維層(Nerve Fiber Layer):

神經細胞的軸突在此層匯聚形成視神經，將視覺信號傳遞到中樞神經系統。

3. 神經節細胞層(Ganglion Cell Layer):

此層負責整合來自神經元的信號，轉換成動作電位傳遞給視神經。

4. 內叢狀層(Inner Plexiform Layer):

此層包含雙極細胞的軸突與無軸突細胞的樹突，將感光細胞輸入的信號傳遞給神經節細胞。

5. 內核層(Inner Nuclear Layer):

此層包含雙極細胞、水平細胞與無軸突細胞的細胞核，幫助調節與傳遞感光信號。

6. 外叢狀層(Outer Plexiform Layer):

感光細胞與雙極細胞、水平細胞之突觸在此層形成。

7. 外核層(Outer Nuclear Layer):

此層包含光感受器細胞的細胞核，是視覺信號產生的起始點。

8. 外界膜(External Limiting Membrane):

由膠質細胞所構成，為感光細胞的細胞核與內節(Inner Segment)之間的屏障，此層可以幫助維持視網膜結構。

9. 感光層(Photoreceptor Layer):

此層由感光細胞組成，是視覺訊號轉導(Phototransduction)的起點。

10. 視網膜色素上皮(Retinal Pigment Epithelium; RPE):

由單層上皮細胞(Epithelial Cells)構成，具有吸收散射光、提供代謝支持與再生視紫質的功能，是感光細胞代謝的關鍵結構。



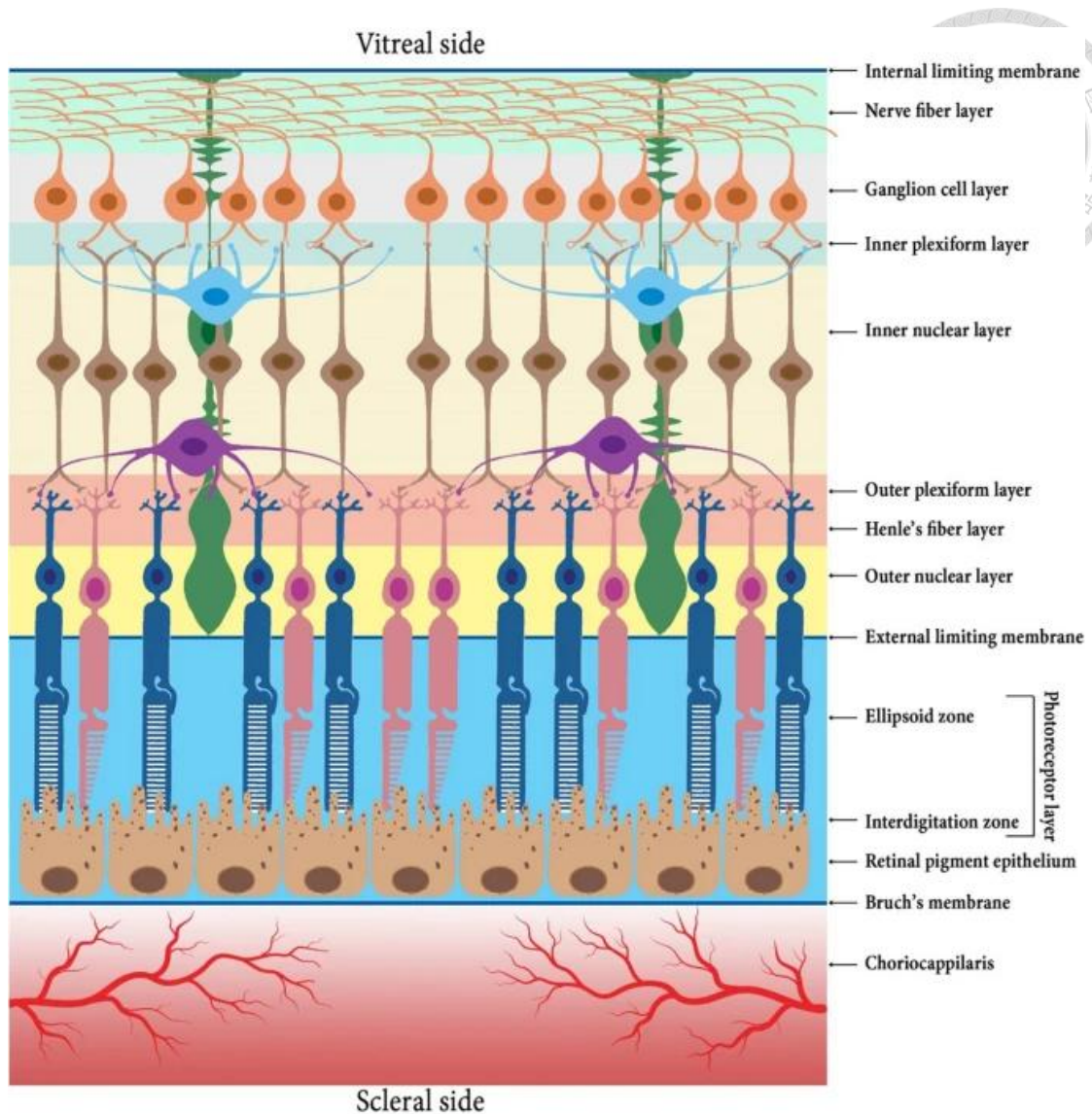


圖 2.3 視網膜十層構造圖[7]

在視網膜中的感光細胞主要包括以下兩種如圖 2.4 所示[8, 9]:

- (1) 桿狀細胞(Rods): 桿狀細胞對微弱光線較為敏感負責暗視覺(Scotopic Vision)。在顏色的分辨上表現有限，但其在光線微弱的環境下可提供基本的視覺感知。
- (2) 錐狀細胞(Cones): 錐狀細胞對亮光較為敏感負責為日間視覺(Photopic Vision)。對顏色的分辨較為敏感其中又可以細分為短波長(S-cones)、中波長(M-cones)與長波長(L-cones)三種類型，分別負責藍、綠與紅色光的辨識。

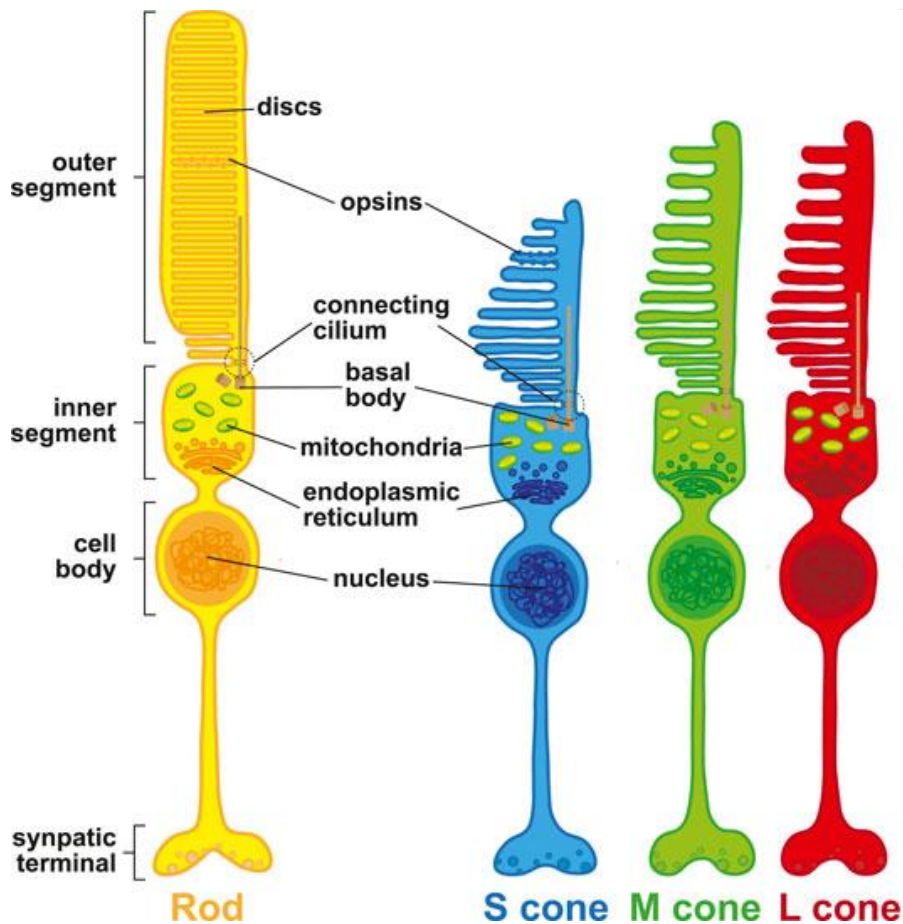


圖 2.4 感光細胞示意圖[8]

2.2.2 黃斑部與中央凹之功能

黃斑部為一個直徑約為 5~5.5 mm 之區域，其中在中央又可以細分出一個約 0.35 mm 之區域稱之為中央凹(Fovea)，在這個區域中有著密度極高的視錐細胞，使得中央凹為人眼視覺解析度(Visual Resolution)的地方，視覺解析度為在兩個人眼能分辨出是兩個點的最小距離，因此人類在閱讀或看東西時常需要轉頭或移動眼睛也是為了能將視物對焦在中央凹上[10—12]。

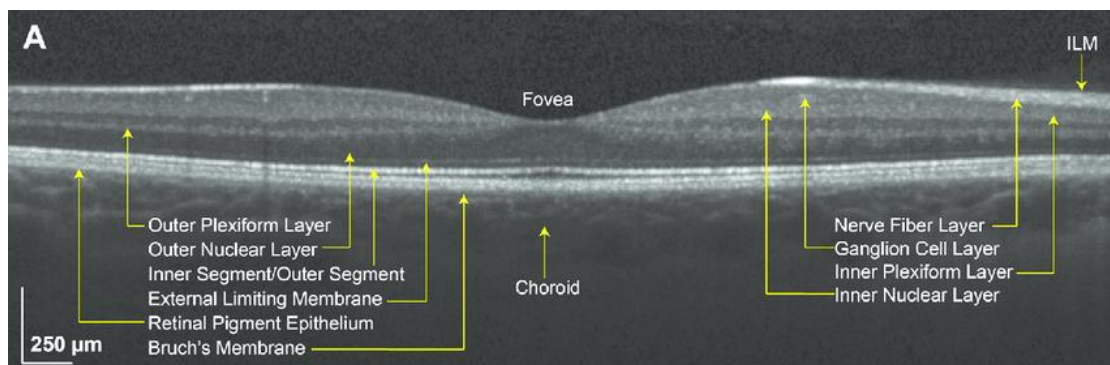


圖 2.5 黃斑部之 OCT 影像[13]



2.3 視網膜疾病

在本節中將會介紹黃斑前膜及在本實驗中是如何定義黃斑前膜之亞健康樣本。

2.3.1 黃斑前膜

黃斑前膜(Epiretinal Membrane; ERM)是由纖維細胞所構成的半透明膜狀結構，形成於內界膜之上，對黃斑部造成拉扯使之變形與導致視力下降，在臨床上 OCT 可以有效幫助 ERM 的診斷[14]。

2.3.2 黃斑前膜之亞健康定義

在本實驗中試圖將健康的視網膜樣本及 ERM 亞健康樣本區分，因此需要在此先定義何為 ERM 亞健康樣本，在所有樣本中兩眼皆未發病則將兩眼皆定義為健康樣本，如圖 2.6 所示；若一隻眼睛未發病而另一隻眼睛確診為 ERM，則將該發病眼睛定義為 ERM 亞健康樣本，如圖 2.7 所示。

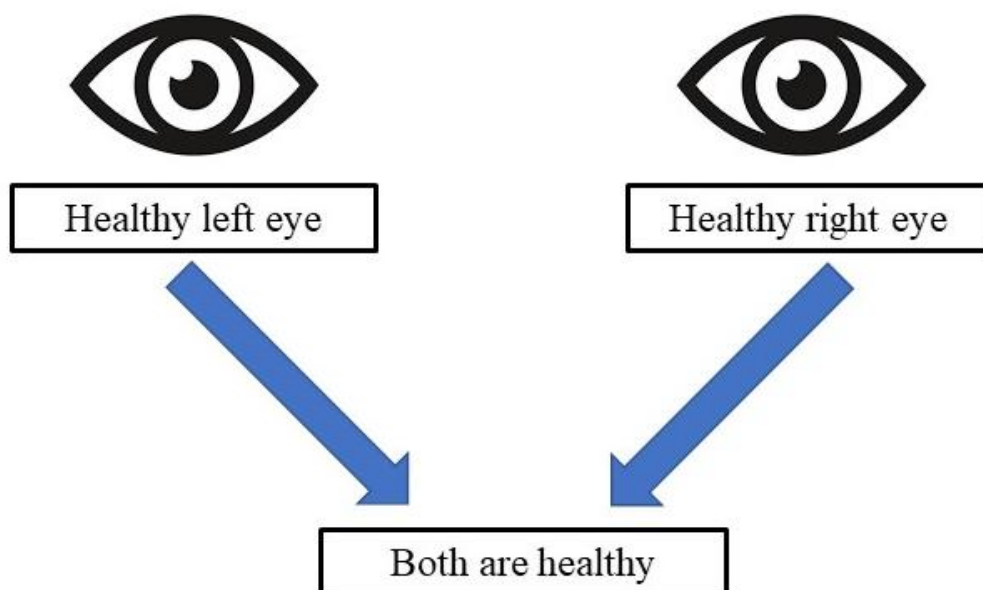


圖 2.6 健康樣本示意圖

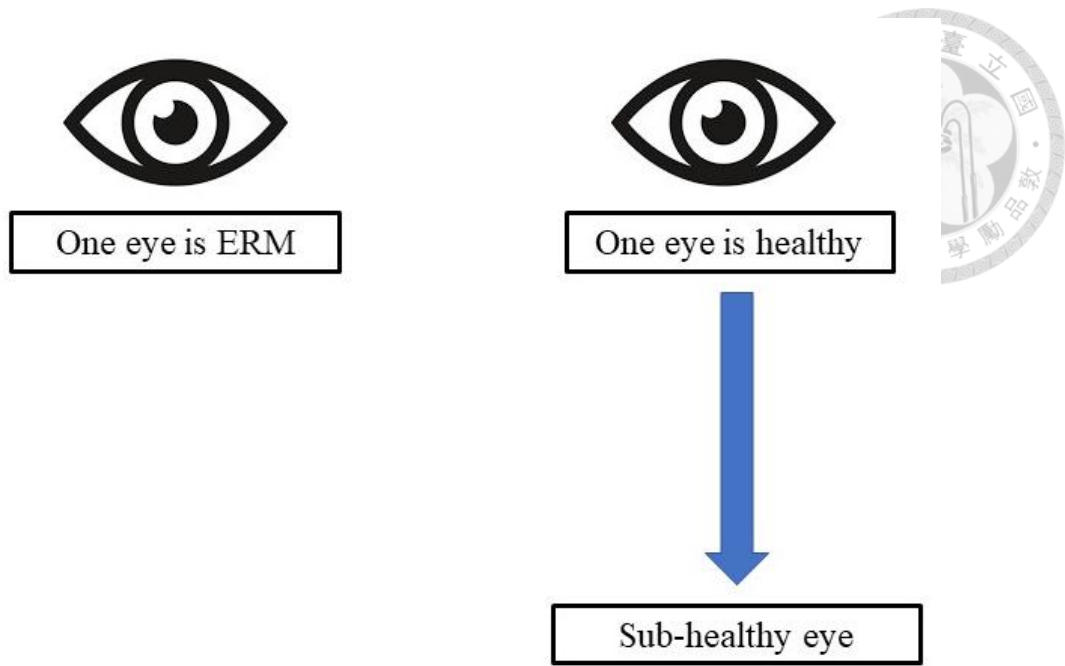


圖 2.7 亞健康樣本示意圖

Chapter 3 研究方法及資料集



在本章中將會對本實驗所使用的研究方法及資料集的收集和對於收集到的影像做了哪些前處理做介紹。

3.1 卷積神經網絡基本架構及殘差網絡架構介紹

卷積神經網絡(Convolutional Neural Network; CNN)是一種目前被廣泛應用在影像分類任務上的深度模型之一，期最早的發展可以追溯到 1989 年所推出的 LeNet-5，此模型成功的達成手寫數字影像的辨識。LeNet-5 的架構包含了輸入層(Input Layer)、卷積層(Convolution Layer)、池化層(Pooling Layer)、全連接層(Fully Connected Layer)與輸出層(Output Layer)，當時所使用的激活函數(Activation function)是 tanh function，且其整體參數量約為 60000 個[15]。LeNet-5 的出現為 CNN 的基本架構奠定了基礎。

在 2012 年時首次有深度學習模型在 ImageNet 的大規模視覺辨識挑戰賽(The ImageNet Large Scale Visual Recognition Challenge; ILSVRC)中取得冠軍，此即為 AlexNet 其參數量約為六千萬(60M)個，且使用了許多現在仍常被使用的方法像是使用 ReLU function (Rectified Linear Unit)來做為激活函數、Dropout、資料擴充(Data Augmentation)等技巧並透過 softmax 函數將模型輸出轉換為機率分佈以進行分類，還有利用 GPU 來進行運算以提高效率[16]。

在此之後大家開始追求更深的模型設計，但越深的模型就越容易產生梯度消失與退化問題，為了解決這些問題 Kaiming He 等人在 2015 年提出了 ResNet 架構，利用殘差學習(Residual Learning)與捷徑連接(Shortcut Connections)的設計，使得超深層的神經網絡結構得以實現。以 ResNet-50 為例該模型的參數量約為兩千五百六十萬個(25.6M)[17]，且該模型是 2015 年的 ILSVRC 冠軍得主。

CNN 的基本架構通常可以細分為以下這六個，接下來將逐一介紹各個部分的功能。

1. 輸入層(Input Layer):

這層負責接收輸入的資料，對於二維影像來說資料的格式通常為高度(Height)×寬度(Width)×通道數(Channels)，為後續層提供訓練來源[18]。

2. 卷積層(Convolution Layer):

這層會透過卷積核(Kernel)在影像上滑動並做內積運算，以此還獲得局部特徵，如圖 3.1 所示，且在多層堆疊後可以獲取更高層更抽象的特徵[2, 5]。

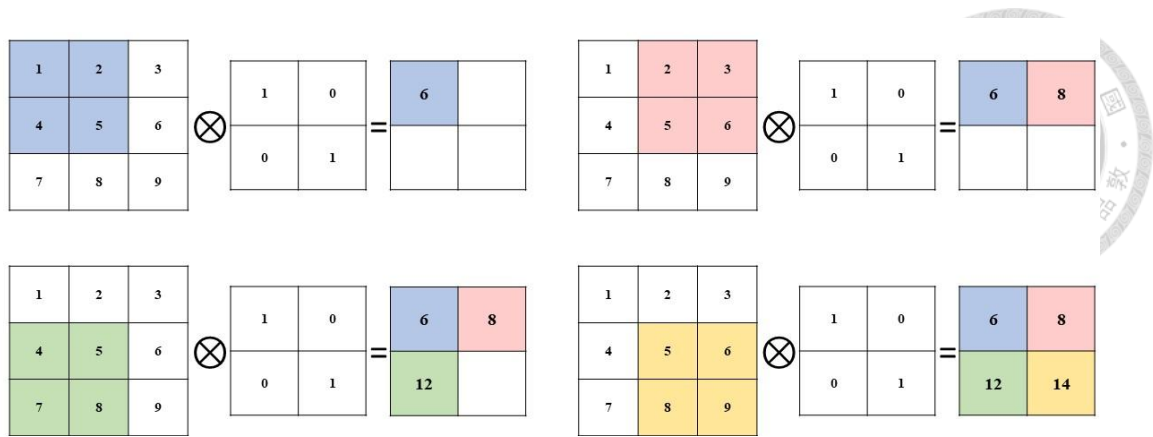


圖 3.1 卷積層示意圖[32]

3. 激活函數(Activation Function):

經過卷積層後的結果為線性輸出，因此透過非線性激活函數來賦予模型學習非線性問題的能力，現在較常使用的激活函數為 ReLU function，可以有效的減少梯度消失的問題[6]。

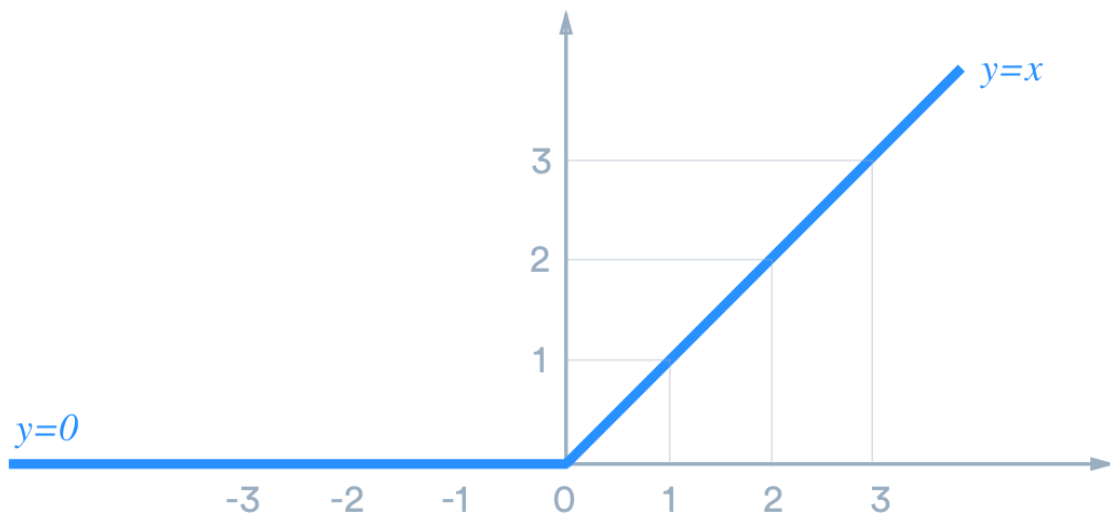


圖 3.2 ReLU function[21]

4. 池化層(Pooling Layer)

這層通常用來降低特徵的空間維度，以此來減少計算成本，同時讓模型構專注於特徵是否存在而非其精確位置，使模型不會變得過度敏感，最常見的為最大池化(Max Pooling)和平均池化(Average Pooling)[15]。

5. 全連接層(Fully Connected Layer):

這層將前面卷積層與池化層所獲得的特徵進行整合送到分類器，此層的每個神經元與前一層的所有神經元相連。

6. 輸出層(Output Layer):

輸出層會因為任務而有所不同，在分類任務中現在通常會使用 Softmax function 將輸出轉為機率分布。

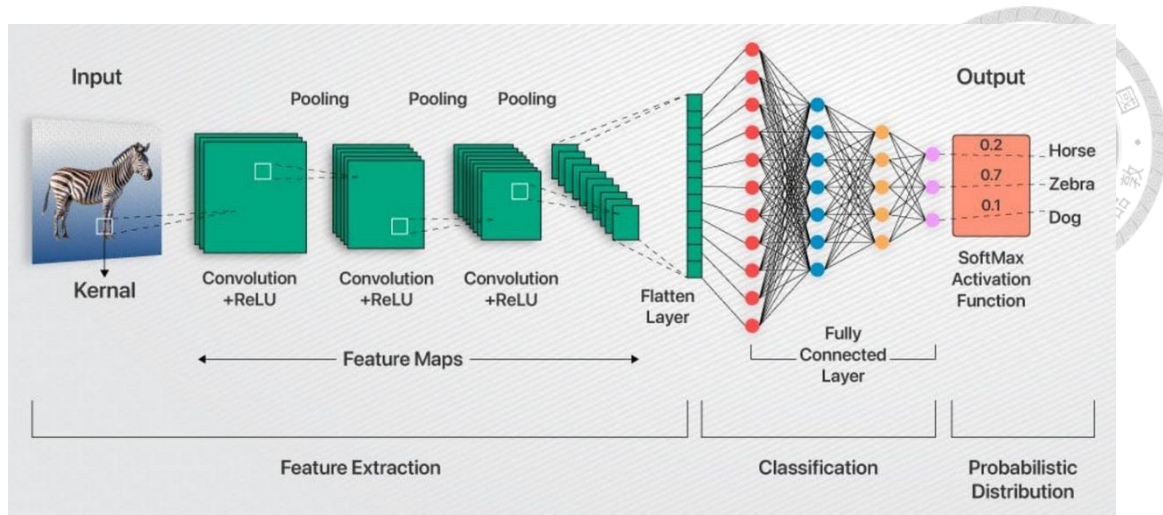


圖 3.3 卷積神經網絡基本架構圖[22]

ResNet 架構與標準的 CNN 架構並沒有太大的區別，最大的不同就在於 ResNet 加入了殘差學習(Residual Learning)的概念，殘差學習的出現解決了度消失的問題，使的超大超深的模型得以實現。

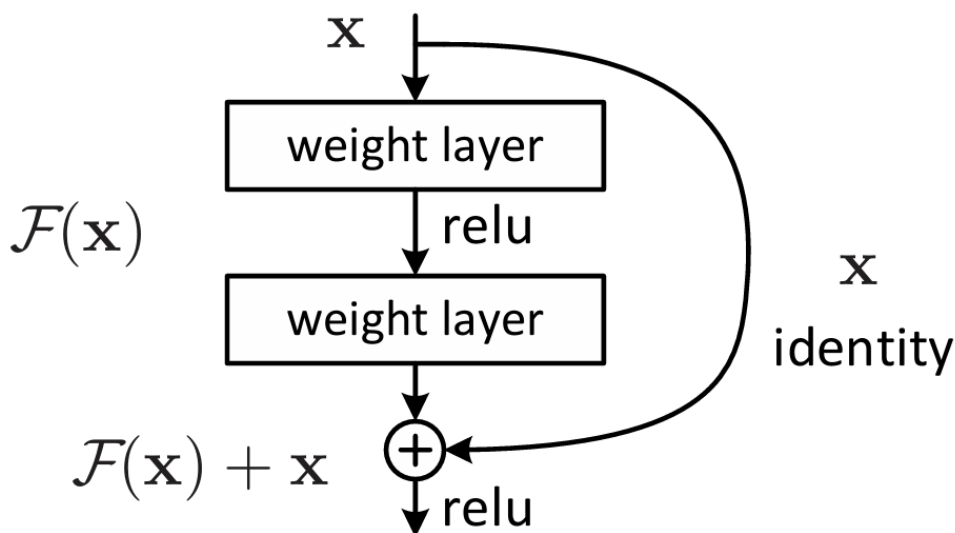


圖 3.4 殘差學習示意圖[17]

在 ResNet 中假設 x 為輸入的資訊， $F(x)$ 表示在這幾層中學習到的殘差，而最終學習到的結果為 $H(x)$ 可以表示為(3.1)。

$$H(x) = F(x) + x \quad (3.1)$$

將(3.1)移項後可以得到殘差 $F(x)$ 為

$$F(x) = H(x) - x \quad (3.2)$$

這樣的改動讓原本模型需要學習「整體要做到甚麼」也就是 $H(x)$ ，改成學習「要改哪些就好」也就是 $F(x)$ 。舉例來說假設輸入 x 為 4.2，經過第一層卷積層後輸出 $H_1(x)$ 為 4.3，根據(3.2)可以得出第一層的殘差 $F_1(x)$ 為 0.1 如(3.3)所示

$$F_1(x) = H_1(x) - x = 4.3 - 4.2 = 0.1 \quad (3.2)$$

假設再經過第二層卷積層後輸出為 $H_2(x)$ 為 4.4，同樣可以計算出第二層與最開始輸入值的殘差 $F_2(x)$ 為 0.2，此時一般網絡與 ResNet 的變化率分別為(3.3)、(3.4)。

$$\frac{H_2(x) - H_1(x)}{H_1(x)} = 0.023 \quad (3.3)$$

$$\frac{F_2(x) - F_1(x)}{F_1(x)} = 1 \quad (3.4)$$

在(3.3)與(3.4)中可以看出殘差學習會比一般的學習容易，這也就是為什麼 ResNet 較容易訓練[23]。

3.2 自監督學習

在本節將會介紹自監督學習與傳統的監督學習有什麼不同，並介紹本實驗所使用的 SimCLR 演算法。

3.2.1 自監督學習與監督學習之介紹與比較

不論是在監督學習(Supervised Learning)還是自監督學習(Self-supervised Learning; SSL)中模型學習的方式都是一樣的，都是透過正向傳播(Forward Propagation)來產生預測結果，並利用反向傳播(Back Propagation)來根據損失函數(Loss Function)的結果更新模型參數，進而逐步最佳化模型的表現。

正向傳播的過程就如同 3.1 節所介紹的，在影像輸入進 CNN 後經過各層的運算後最後透過 Softmax 運算得到一個機率分布。

損失函數就是用來評估這個模型學得如何的指標，當 loss 越小時通常也表示模型學的越好，準確率越高，如(3.5)所示 y 為實際值； \hat{y} 為預測值。

$$Loss = y - \hat{y} \quad (3.5)$$

在分類任務中較常使用的三個損失函數為均方誤差(Mean Square Error; MSE)、平均絕對值誤差(Mean Absolute Error; MAE)及交叉熵(Cross Entropy)[24, 25]。

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3.6)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3.7)$$

在 MSE 及 MAE 中 n 表示有 n 筆資料； y 為實際值； \hat{y} 為預測值。

$$Cross\ Entropy = \sum_{c=1}^C \sum_{i=1}^n -y_{c,i} \log_2(p_{c,i}) \quad (3.8)$$

在交叉熵中 C 為總共有幾類； n 表示有 n 筆資料； $y_{c,i}$ 為第 i 筆資料實際屬於第 c

類； $p_{c,i}$ 為第 i 筆資料預測屬於第 c 類。

反向傳播(Back Propagation)是透過損失函數計算模型預測與實際目標之間的差距，再利用連鎖律(Chain Rule)計算出各層的梯度(Gradient)從輸出層反向傳回輸入層，最後使用不同的優化器(Optimizer)現多使用 Adam 進行參數更新，使模型收斂找到最佳解[26]。

在監督學習中，訓練模型需要大量有標註的資料，每個樣本都有對應的標籤(Label)，用來提供模型標準答案，透過不斷學習這些樣本的對應關係以及損失函數的最佳化來提升模型的準確率，但是這些樣本的標註卻需要耗費大量的時間。

在自監督訓練中則使用不同的方式產生偽標籤(Pseudo Label)，像是旋轉、裁切等方式，透過這些方式來模型可以在不需要標註資料的情況下學習到資料的特徵，在後續微調(Fine Tuning)階段應用至下游任務，達到不錯的表現。

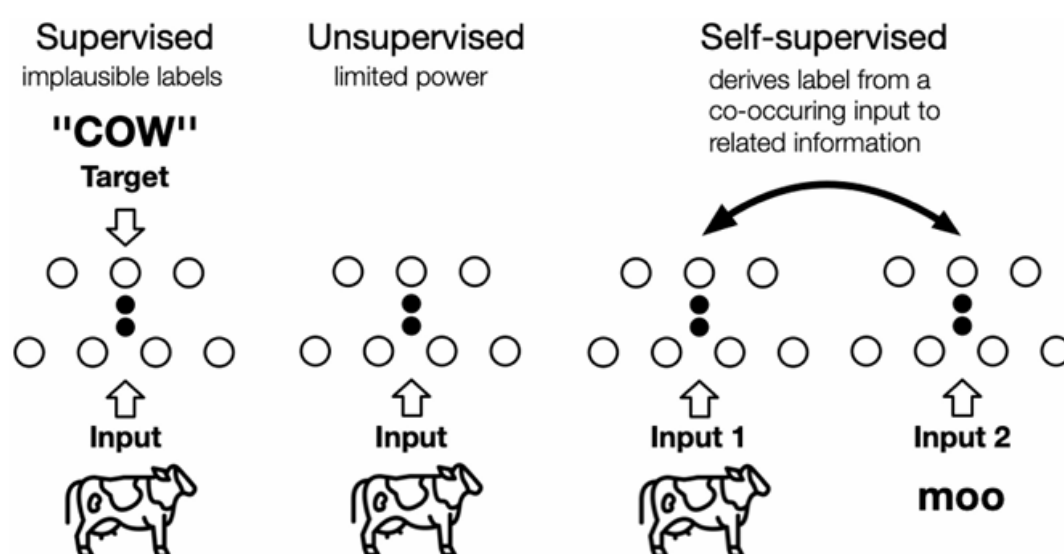
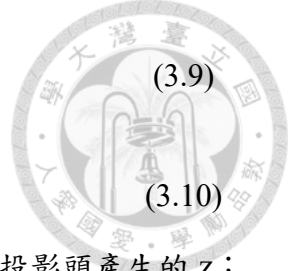


圖 3.5 監督學習與自監督學習示意圖[27]

3.2.2 SimCLR 流程及應用

本實驗所使用的自監督學習演算法為 SimCLR(A Simple Framework for Contrastive Learning of Visual Representations)，將影像輸入後會對所有生成不同的兩個增強樣本，彼此之間互為正樣本(Positive Pair)，用此方式來產生偽標籤，再透過編碼器(Encoder)來提取特徵向量 h ，這邊的編碼器通常使用不同大小的 ResNet，模型因需求而做更改，再利用一個投影頭(Projection Head)將 h 投影成 z 來做損失函數的計算，因為作者發現雖然特徵向量 h 適合做後續的分類但如果直接用來做損失函數的計算反而會導致結果變差因此才投影成 z 再做計算，SimCLR 中所使用的損失函數如(3.9)所示。



$$l_{i,j} = -\log \frac{\exp\left[\frac{\text{sim}(z_i, z_j)}{\tau}\right]}{\sum_{k=1}^{2N} \mathbf{1}_{[k \neq i]} \exp\left[\frac{\text{sim}(z_i, z_k)}{\tau}\right]} \quad (3.9)$$

$$\text{sim}(u,v) \equiv \frac{u^\top v}{\|u\| \|v\|} \quad (3.10)$$

其中 $l_{i,j}$ 為第 i 個樣本與其正樣本 j 的 loss； z_i, z_j 為經過編碼器和投影頭產生的 z ； $\text{sim}(u,v)$ 為餘弦相似性(Cosine Similarity)； τ 為溫度參數(Temperature Parameter)； k 為其他所有可能拿來做比較的樣本[28]。

Algorithm 1 SimCLR's main learning algorithm.

```

input: batch size  $N$ , constant  $\tau$ , structure of  $f, g, \mathcal{T}$ .
for sampled minibatch  $\{\mathbf{x}_k\}_{k=1}^N$  do
  for all  $k \in \{1, \dots, N\}$  do
    draw two augmentation functions  $t \sim \mathcal{T}, t' \sim \mathcal{T}$ 
    # the first augmentation
     $\tilde{\mathbf{x}}_{2k-1} = t(\mathbf{x}_k)$ 
     $\mathbf{h}_{2k-1} = f(\tilde{\mathbf{x}}_{2k-1})$  # representation
     $\mathbf{z}_{2k-1} = g(\mathbf{h}_{2k-1})$  # projection
    # the second augmentation
     $\tilde{\mathbf{x}}_{2k} = t'(\mathbf{x}_k)$ 
     $\mathbf{h}_{2k} = f(\tilde{\mathbf{x}}_{2k})$  # representation
     $\mathbf{z}_{2k} = g(\mathbf{h}_{2k})$  # projection
  end for
  for all  $i \in \{1, \dots, 2N\}$  and  $j \in \{1, \dots, 2N\}$  do
     $s_{i,j} = \mathbf{z}_i^\top \mathbf{z}_j / (\|\mathbf{z}_i\| \|\mathbf{z}_j\|)$  # pairwise similarity
  end for
  define  $\ell(i, j)$  as  $\ell(i, j) = -\log \frac{\exp(s_{i,j}/\tau)}{\sum_{k=1}^{2N} \mathbf{1}_{[k \neq i]} \exp(s_{i,k}/\tau)}$ 
   $\mathcal{L} = \frac{1}{2N} \sum_{k=1}^N [\ell(2k-1, 2k) + \ell(2k, 2k-1)]$ 
  update networks  $f$  and  $g$  to minimize  $\mathcal{L}$ 
end for
return encoder network  $f(\cdot)$ , and throw away  $g(\cdot)$ 

```

圖 3.6 SimCLR 訓練流程[28]

3.3 聚類方式

在本節中將介紹本實驗所使用的無監督聚類方式 K-means Clustering 與半監督聚類方式 SDEC (Semi-supervised Deep Embedded Clustering)



3.3.1 無監督聚類

K-means 是一個常見的無監督聚類方式，主要透過最小化樣本與所屬群心 (Centroid) 之間的距離平方和 (Within-cluster Sum of Squares; WCSS) 來進行分群，其目標函數如(3.11)所示[29]。

$$J = \sum_{i=1}^K \sum_{x \in C_i} |x - \mu_i|^2 \quad (3.11)$$

其中 K 為指定的群數； μ_i 為第 i 群之群心； C_i 為第 i 群集所包含之樣本集合。但當輸入的資料開始越來越複雜後，單純使用 K-means 聚類的結果十分受限[30]，因此 J. Xie 等人提出 DEC (Deep Embedding Clustering) 方法，將深度學習模型與 K-means 聚類結合使用，先利用 CNN 來獲得資料的嵌入特徵向量 (Embedding)，再利用這個嵌入特徵向量來做 K-means 聚類[31]，其目標函數如(3.12)所示。

$$L = \text{KL}(P \parallel Q) = \sum_i \sum_j p_{ij} \log \left(\frac{p_{ij}}{q_{ij}} \right) \quad (3.11)$$

$$p_{ij} \equiv \frac{\frac{q_{ij}^2}{\sum_i q_{ij}}}{\sum_j \frac{q_{ij}^2}{\sum_i q_{ij}}} \quad (3.12)$$

其中 i 為第 i 個樣本；j 為第 j 個群集； q_{ij} 預測樣本 i 為第 j 群的機率。

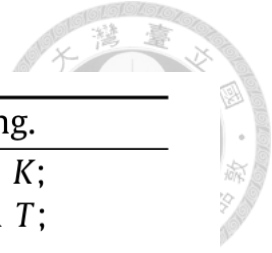
3.3.2 半監督聚類

當資料中存在樣本異質性高、群界模糊或部分樣本可信度不足時，DEC 的結果可能也將受限，因此 Ren 等人提出了 SDEC (Semi-supervised Deep Embedded Clustering) 方法，先將部分資料的標籤 (Label) 告訴模型建立配對資訊 (Pairwise Constraints)，讓模型知道哪些資料應該靠近哪些資料應該遠離，其目標函數為 DEC 之目標函數加上配對損失函數如(3.13)及(3.14)所示。

$$L = \text{KL}(P \parallel Q) + \lambda L_{\text{pairwise}} \quad (3.13)$$

$$L_{\text{pairwise}} = \frac{1}{n} \sum_{i=1}^n \sum_{k=1}^n a_{ik} |z_i - z_k|^2 \quad (3.14)$$

其中 $\text{KL}(P \parallel Q)$ 為 DEC 之目標函數； λ 為 L_{pairwise} 之權重；n 為總樣本數； a_{ik} 為 i 樣本與 k 樣本之配對限制指示參數 (Pairwise Constraint Indicator)，1 表示應該同一群、-1 表示應該不同群、0 則表示兩個樣本至少其中一個的標籤並未告訴模型； z_i 、 z_k 為樣本之嵌入特徵向量。



Algorithm 1 Semi-supervised deep embedded clustering.

Require: Data set X ; coefficient λ ; number of clusters K ;
pairwise constraints matrix A ; update interval T ;
stopping threshold $tol\%$.

Ensure: Cluster assignments $\{y_i\}_{i=1}^n$; cluster centers $\{\mu_i\}_{i=1}^K$; deep
mapping f_θ .

- 1: **Step 1** \rightarrow **Initialization with SAE**
 - 2: Pretrain SAE and obtain K initial centers $\{\mu_j\}_{j=1}^K$ and cluster
assignments $\{y_i\}_{i=1}^n$ by running k -means in the latent space Z .
 - 3: **Step 2** \rightarrow **Clustering with pairwise constraints**
 - 4: **for** $iter \in \{0, 1, \dots, MAXITER\}$ **do**
 - 5: Choose a batch of samples $S \subset X$.
 - 6: **if** $iter \% T == 0$ **then**
 - 7: $z_i \leftarrow f_\theta(x_i), \forall x_i \in X$.
 - 8: Compute all q_{ij} values according to Eq. (1).
 - 9: Compute all p_{ij} values according to Eq. (2).
 - 10: Save old assignments: $y_{oldi} \leftarrow y_i$.
 - 11: Update label assignments: $y_i \leftarrow \arg \max_j q_{ij}$.
 - 12: **if** $(\sum_{i=1}^n y_{oldi} \neq y_i) / n < tol\%$ **then**
 - 13: Stop training
 - 14: **end if**
 - 15: **end if**
 - 16: Update θ and $\{\mu_j\}_{j=1}^K$ via Eqs. (5) and (6).
 - 17: **end for**
-

圖 3.7 SDEC 訓練流程[29]

在圖 3.7 中可以看到首先會先訓練一個堆疊式自編碼器(Stacked Auto Encoder ; SAE) 來獲得資料的嵌入特徵向量，但在本實驗中將會用不同預訓練方式所訓練之模型來獲得資料的嵌入特徵向量，因此不會重新訓練一個堆疊式自編碼器，獲得嵌入特徵向量後首先會對其做 K -means 聚類以此來獲得初始的群心(Cluster Centers)，接下來根據部分資料的標籤獲得的配對限制指示參數搭配(3.13)讓模型不斷迭代，直到模型預測的資料標籤變化小於設定的更新間隔(Update Interval; T)或達到最大迭代次數。此演算法可以讓模型在對資料進行聚類時同時考慮資料內在的潛在結構及人工標註的指引，進而得到較符合實際情況的預測標籤。



3.4 資料集與前處理

在本節中將會說明本實驗的資料量集來源，以及對原始資料做了哪些前處理。

3.4.1 資料集之來源及數量

本實驗所使用的 OCT 影像為 2012 年 1 月到 2019 年 12 月期間在瑞光眼科診所接受黃斑部 OCT 檢查之年滿 20 歲病患視網膜影像，所以影像皆使用 Heidelberg Spectralis OCT 拍攝，該儀器由淺至深總構拍攝 25 張影像，本實驗只使用第 11 張影像來訓練模型。

本實驗所用到的 OCT 影像總共有 19464 張，這裡面包含所有健康、亞健康及黃斑部結構異常之資料，又可以將這些資料分成以下四個資料集：

1. Healthy vs Unhealthy 資料集：

在這個資料集中共有 9732 張 OCT 影像，其中有 4866 張 OCT 影像來自 2288 位健康受試者；4866 張 OCT 影像來自 2043 位黃斑部結構異常(包含 ERM)之病患。

2. ERM vs Others diseases 資料集：

在這個資料集中共有 3748 張 OCT 影像，其中有 1874 張 OCT 影像來自 828 位患有 ERM 之病患；1874 張 OCT 影像來自 913 位黃斑部結構異常(不包含 ERM)之病患，如糖尿病黃斑水腫(Diabetic Macular Edema; DME)、老年性黃斑部病變(Age-related Macular Degeneration; AMD)等等。

3. ERM vs Healthy 資料集：

在這個資料集中共有 1142 張 OCT 影像，其中有 606 張 OCT 影像來自 556 位健康受試者；536 張 OCT 影像來自 254 位患有 ERM 之病患。

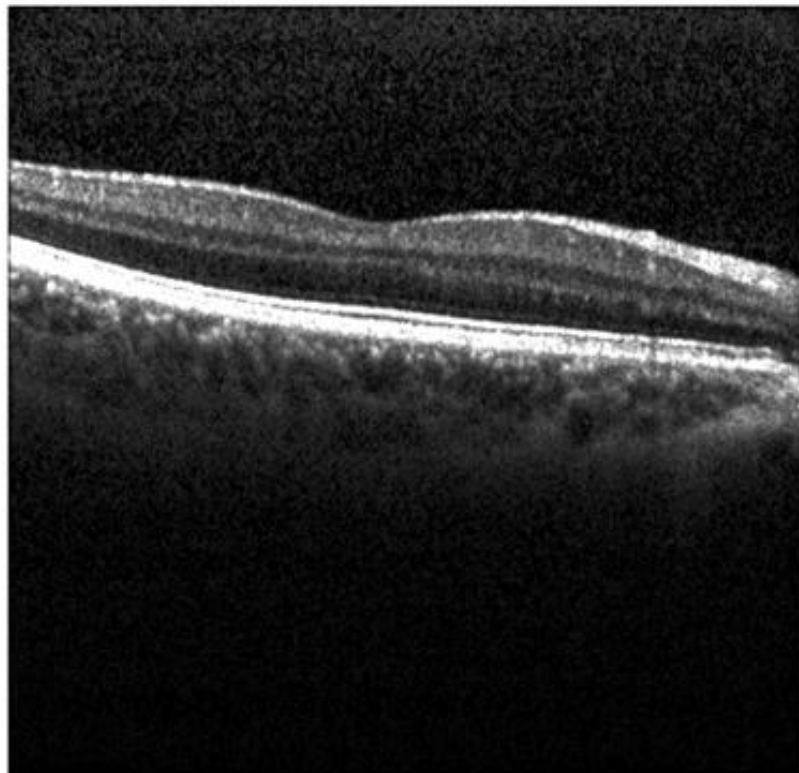
4. Healthy vs Sub-healthy(Female)資料集：

在這個資料集中共有 764 張 OCT 影像，其中有 400 張 OCT 影像來自 303 位健康受試者；364 張 OCT 影像來自 162 位患有 ERM 之病患，且在此資料集中的病患皆為女性。

3.4.2 影像之前處理

Heidelberg Spectralis OCT 儀器所拍攝之原始影像之尺寸為 512×496 (Pixel)，軸向與橫向解析度分別為 3.9 μm 與 11 μm ，將上下多於黑色部分裁減掉後為 512×256 (Pixel)如圖 3.8、圖 3.9 所示。

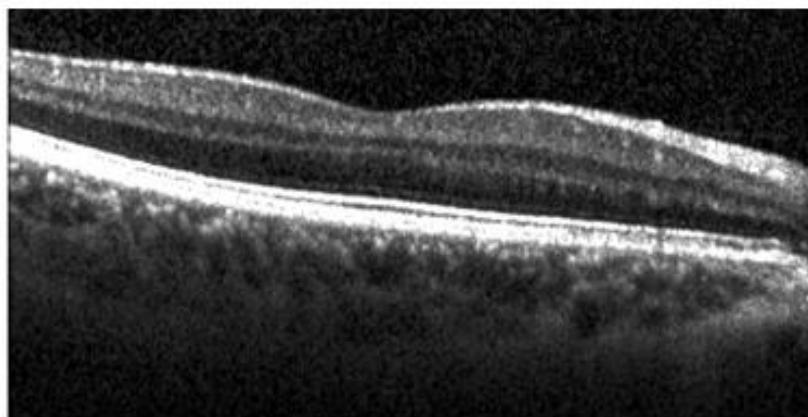
496 pixels



512 pixels

圖 3.8 視網膜 OCT 原始影像

256 pixels



512 pixels

圖 3.9 視網膜 OCT 裁切後之影像

Chapter 4 健康、亞健康與不健康眼睛之分類



在本章中將會介紹本實驗所使用之深度學習模型及其參數設定，並對展示健康與不健康、黃斑前膜與其他疾病、健康與黃斑前膜、健康與亞健康四項分類任務之結果與分析。

4.1 實驗設計

在本節中將會說明本實驗訓練所使用的深度學習模型架構、其使用的各項超參數 (Hyperparameters)、預訓練方式、深度學習模型訓練時的設定。

4.1.1 模型架構及預訓練方式之介紹

在本實驗中所使用的深度學習模型皆為 ResNet-18 模型，架構上的更改只有將輸入層接收的通道數從三改為一，也就是從彩色影像 (三通道) 的輸入改為灰階影像 (單通道) 的輸入，較符合現有的 OCT 影像格式，ResNet-18 的參數量約為一千一百萬個 (11M)，其架構如下表所示。

表 4.1 本實驗中所使用之模型架構[17]

層名稱	Conv/FC 結構
Conv1	7×7, 64, stride 2
MaxPool	3×3 max pool, stride 2
Conv2_x	$\begin{bmatrix} 3 \times 3, & 64 \\ 3 \times 3, & 64 \end{bmatrix} \times 2$ (Residual)
Conv3_x	$\begin{bmatrix} 3 \times 3, & 128 \\ 3 \times 3, & 128 \end{bmatrix} \times 2$ (Residual)
Conv4_x	$\begin{bmatrix} 3 \times 3, & 256 \\ 3 \times 3, & 256 \end{bmatrix} \times 2$ (Residual)
Conv5_x	$\begin{bmatrix} 3 \times 3, & 512 \\ 3 \times 3, & 512 \end{bmatrix} \times 2$ (Residual)
Average Pool	Adaptive Average Pool2d (512 to 512)
Fully Connected Layer	Dropout (p=0.5) Linear (512 to 256) Dropout (p=0.5) Linear (256 to 64) Dropout (p=0.5) Linear (64 to 2) Softmax (1)

在本實驗中總共使用三種預訓練方式，分別為使用 ImageNet 預訓練的 ResNet-18 模型，在此預訓練方式中使用在 ImageNet 圖庫中約 1280 萬張共 1000 個類別的影像來進行監督式學習，後續將稱此模型為 ImageNet Pre-train Model；第二種為隨機初始權重的 ResNet-18 模型使用總共 19464 張黃斑部之 OCT 影像來進行 SimCLR 之自監督學習所訓練的 ResNet-18 模型，後續將稱此模型為 SimCLR Pre-train Model；第三種為在先前研究中所使用的 ResNet-18 模型，先使用 ImageNet 預訓練的 ResNet-18 模型在使用約 4000 張黃斑部之 OCT 影像進行微調(Fine Tune)在男女分類任務上，後續將稱此模型為 OCT Pre-train Model [33]。在本實驗進行的四個分類任務中針對以下三類任務:健康與不健康之分類、黃斑前膜與其他疾病之分類、健康與黃斑前膜之分類，將比較 ImageNet Pre-train Model 與 SimCLR Pre-train Model 之表現，觀察並分析自監督學習於醫學影像分類任務中所帶來之效益與潛力。在健康與黃斑前膜之分類中將比較三個不同預訓練模型之結果，觀察在架構相同的情況下使用不同預訓練方式是否能更有效的辨別出亞健康影像。

從圖 4.1 與圖 4.2 可觀察到，在 SimCLR 自監督學習的預訓練過程中，驗證集的準確率最高可達 92.84%。同時，訓練與驗證的準確率曲線在後期皆趨於平穩，未再出現顯著變動，Loss 曲線亦呈現相同趨勢，顯示模型已達收斂。因此可確認預訓練階段已順利完成，後續任務將以此預訓練模型進行。

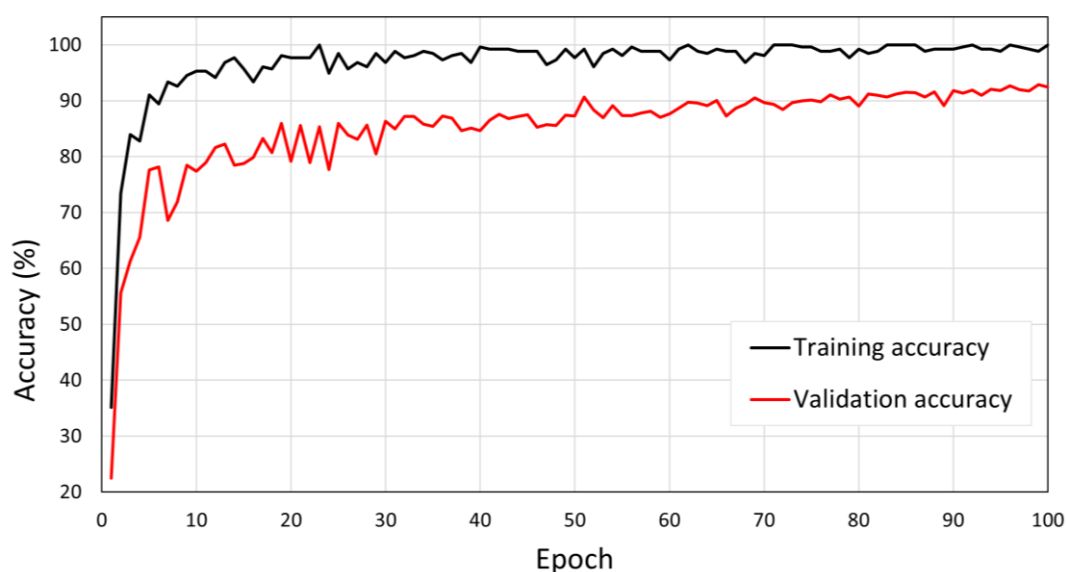


圖 4.1 在 SimCLR 自監督學習預訓練中訓練集與驗證集對影像增強產生偽標籤對分類之準確率

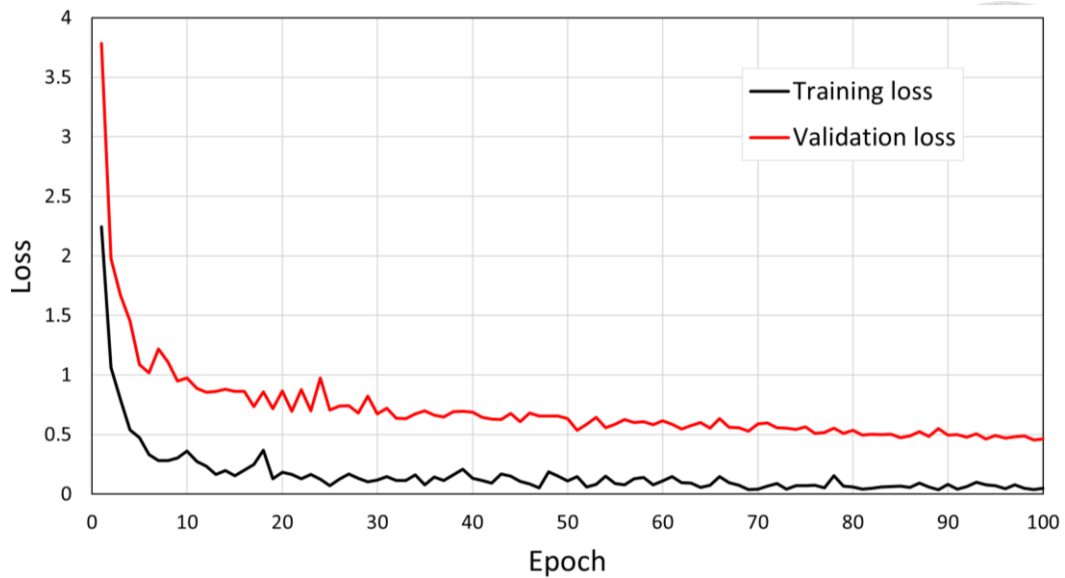


圖 4.2 在 SimCLR 自監督學習預訓練中訓練集與驗證集對影像增強產生偽標籤對分類之 Loss

4.1.2 訓練設定及評估指標

在本章的實驗中均將資料集以訓練 (Training)使用 80%、驗證 (Validation)使用 10%、測試 (Testing)使用 10%的方式進行切割，並採用交叉驗證 (Cross Validation) 使在訓練、驗證、測試中可以使用到多種不同的資料集組合。在本實驗中會將資料分為十等分，也就是進行 10-Fold Cross Validation，將資料分為十等份後首先會先選取一份資料做為測試集，然後在剩下的九份中選取一份做為驗證集，最後剩下的八份即為此次訓練中的訓練集，並且測試集與驗證集在十次的訓練中並不會重複選取，也就是每次訓練時所使用的測試集與驗證集皆不同如圖 4.3 所示，最終的測試結果將會以十次的測試準確率及其標準差來進行分析。

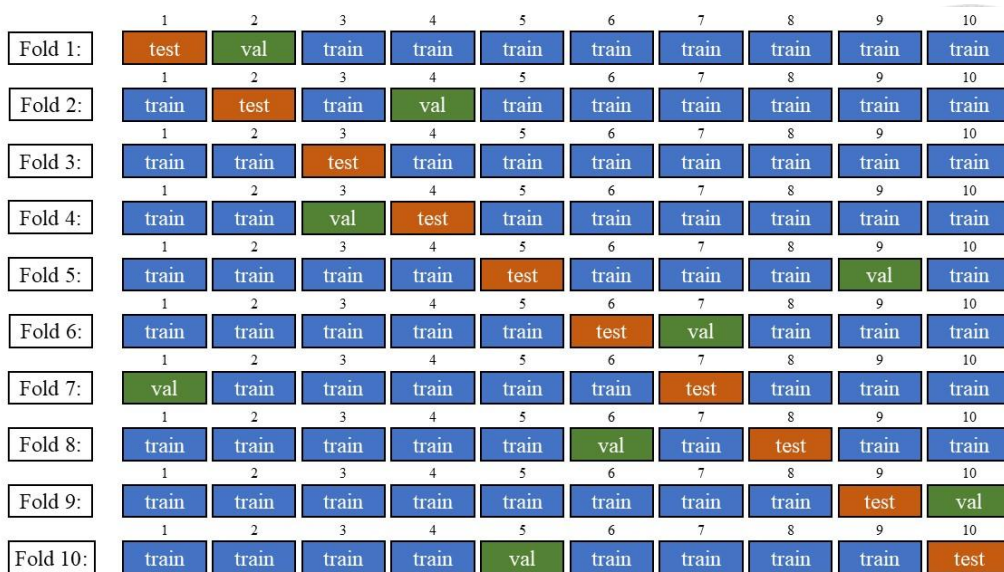


圖 4.3 Cross Validation 示意圖

4.2 分類任務一：健康與不健康之分類

在此任務中將使用「Healthy vs. Unhealthy 資料集」來做模型的訓練與測試，共有 9732 張 OCT 影像，其中有 4866 張 OCT 影像來自 2288 位健康受試者；4866 張 OCT 影像來自 2043 位黃斑部結構異常(包含 ERM)之病患。其中 80%用來做 training (7785 張)；10%用來做 validation (973 張)；10%用來做 testing (974 張)。在本任務中所使用的模型為 ImageNet Pre-train Model 及 SimCLR Pre-train Model，各項超參數(Hyperparameters)設定如表 4.2 所示。

表 4.2 健康與不健康之分類任務中各模型之超參數設定

	ImageNet Pre-train Model	SimCLR Pre-train Model
超參數	Epoch: 70 Batch size: 32 Optimizer: Adam Initial learning rate: 5×10^{-7} Weight decay: 1×10^{-4}	Epoch: 70 Batch size: 32 Optimizer: Adam Initial learning rate: 5×10^{-7} Weight decay: 1×10^{-4}

在表 4.3 中可以看到 ImageNet Pre-train Model 在 10 次交叉驗證訓練後其在測試集中平均準確率為 83.3%，以及平均 0.87%的標準差；在表 4.4 中可以看到 SimCLR Pre-train Model 在 10 次交叉驗證訓練後其在測試集中平均準確率為 80.31%，以及平均 1.06%的標準差。兩者皆能達到 80%以上的準確率，皆能正確的辨別健康與不健康的視網膜 OCT 影像，但可以觀察到 ImageNet Pre-train Model 的表現略優於 SimCLR Pre-train Model。

在表 4.5 中的 Grad-CAM 影像中可以觀察到 ImageNet Pre-train Model 與

SimCLR Pre-train Model 在看健康影像時皆會將重點放在中央凹的位置，但 SimCLR Pre-train Model 除了中央凹外還會將重點放在其他的部分；在表 4.6 中的 Grad-CAM 影像中可以觀察到 ImageNet Pre-train Model 與 SimCLR Pre-train Model 在看不健康影像時皆可以將重點放在有異常的區域。

表 4.3 ImageNet Pre-train Model 在健康眼睛與不健康眼睛(包含 ERM)分類任務中測試集之表現結果

	1	2	3	4	5	6	7	8	9	10	Avg ±SD
Accuracy (%)	81.9	83.8	83.9	84.6	84.3	83.5	83	83.5	82.2	82.3	83.3 ±0.87

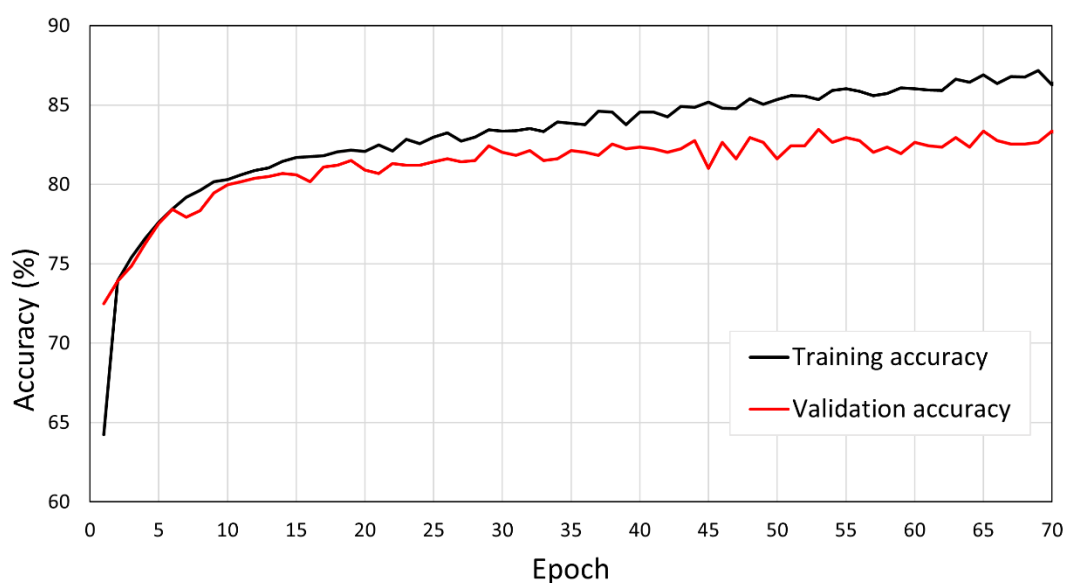


圖 4.4 ImageNet Pre-train Model 在健康眼睛與不健康眼睛(包含 ERM)分類任務中訓練集與驗證集準確率與 Epoch 做圖結果

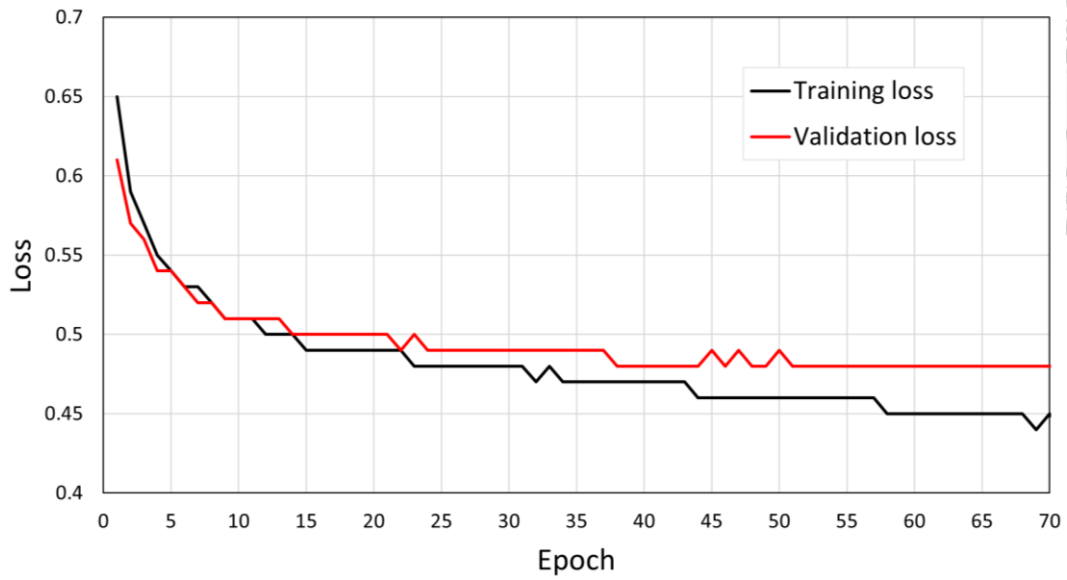


圖 4.5 ImageNet Pre-train Model 在健康眼睛與不健康眼睛(包含 ERM)分類任務中訓練集與驗證集 Loss 與 Epoch 做圖結果

表 4.4 SimCLR Pre-train Model 在健康眼睛與不健康眼睛(包含 ERM)之分類任務中測試集之表現結果

	1	2	3	4	5	6	7	8	9	10	Avg ±SD
Accuracy (%)	79.8	79.8	80.8	81.4	81.9	80	79.3	81.6	80.2	78.3	80.31 ±1.06

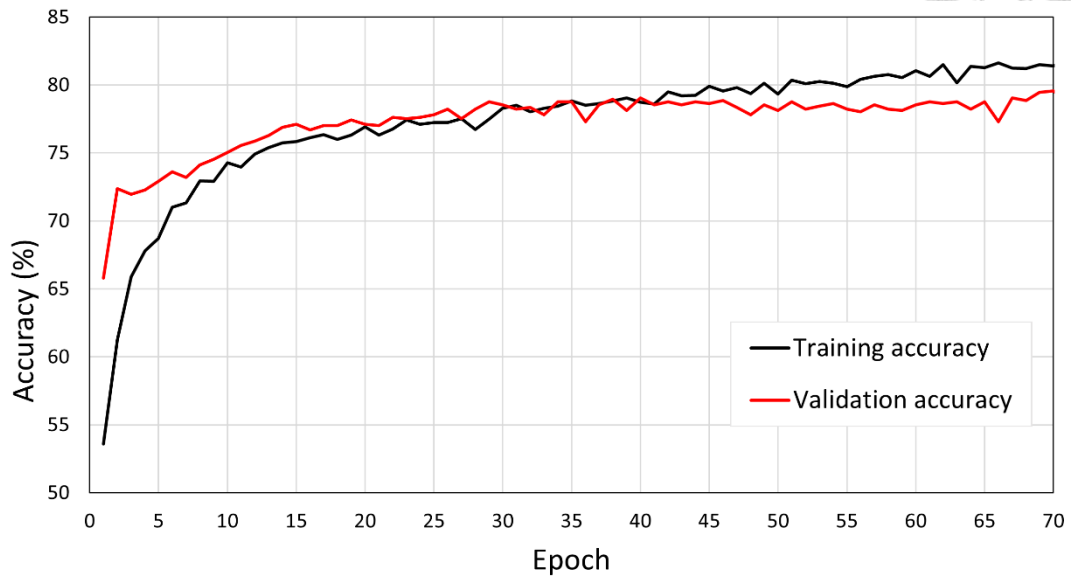


圖 4.6 SimCLR Pre-train Model 在健康眼睛與不健康眼睛(包含 ERM)分類任務中
訓練集與驗證集準確率與 Epoch 做圖結果

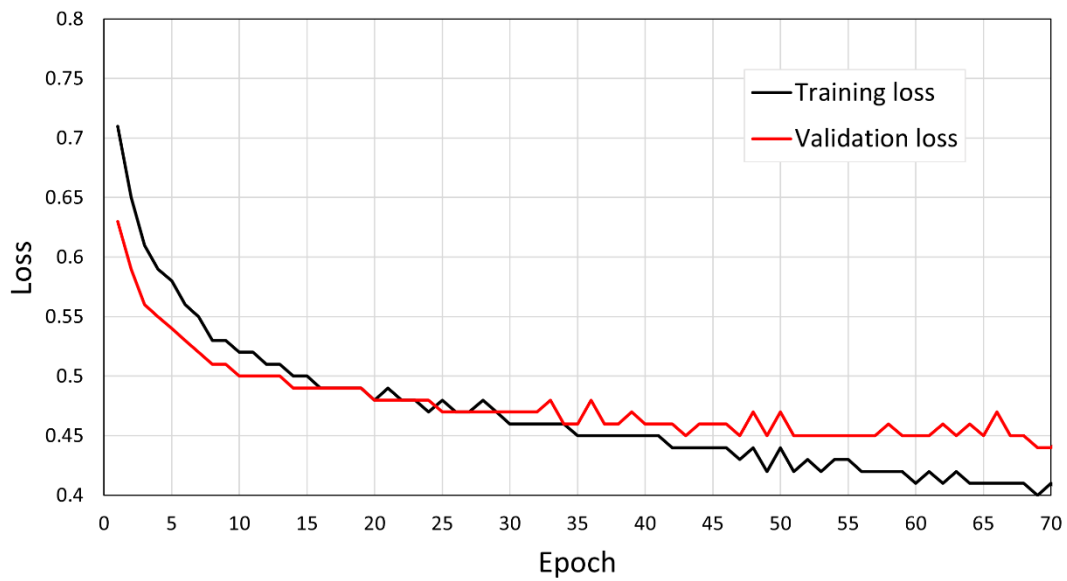


圖 4.7 SimCLR Pre-train Model 在健康眼睛與不健康眼睛(包含 ERM)分類任務中
訓練集與驗證集 Loss 與 Epoch 做圖結果

表 4.5 ImageNet Pre-train Model 及 SimCLR Pre-train Model 經健康與不健康分類
任務訓練後最後一層對健康影像之 Grad-CAM 影像

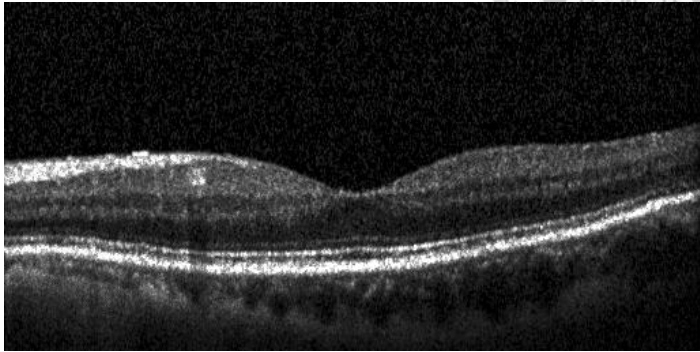
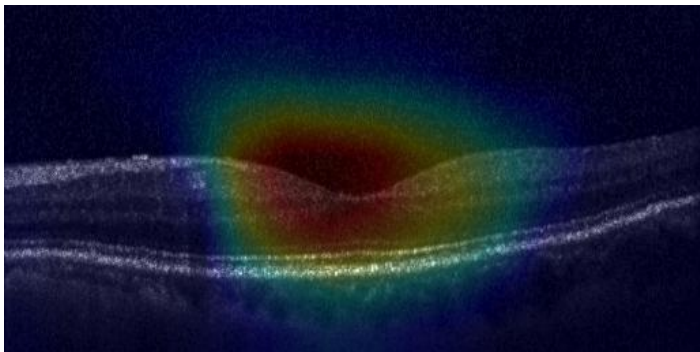
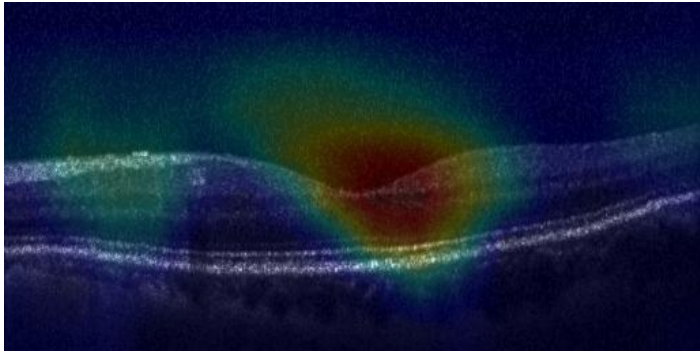
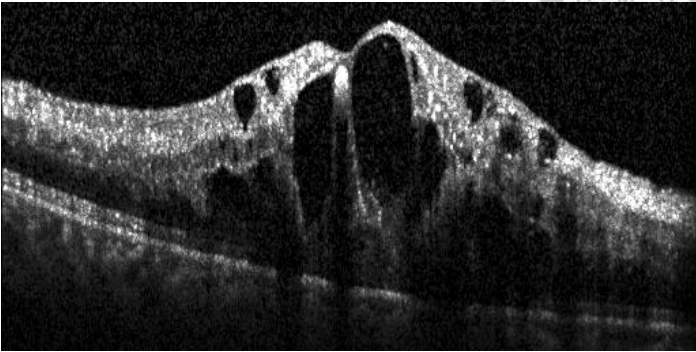
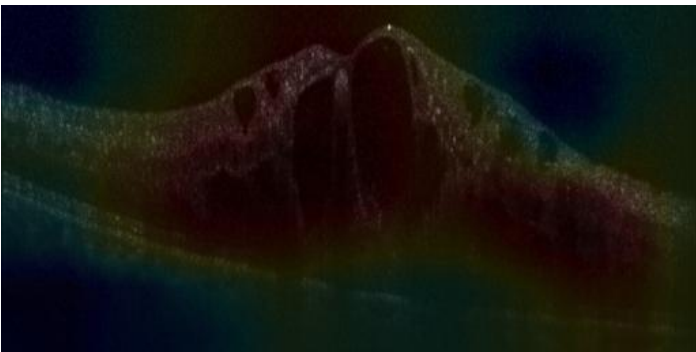
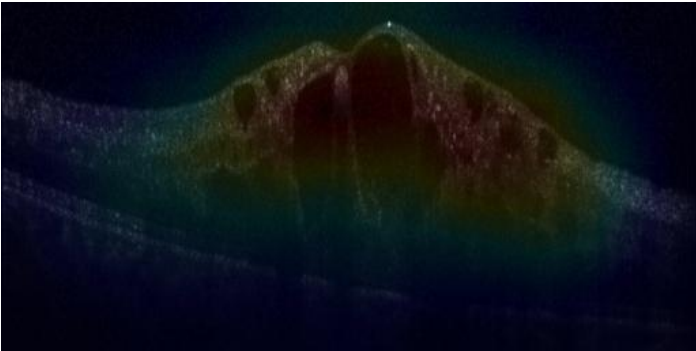
<p>Input Image (Healthy)</p>	
<p>Layer4 Grad-CAM (ImageNet Pre-train Model)</p>	
<p>Layer4 Grad-CAM (SimCLR Pre-train Model)</p>	

表 4.6 ImageNet Pre-train Model 及 SimCLR Pre-train Model 經健康與不健康分類任務訓練後最後一層對不健康影像之 Grad-CAM 影像

<p>Input Image (Unhealthy)</p>	
<p>Layer4 Grad-CAM (ImageNet Pre-train Model)</p>	
<p>Layer4 Grad-CAM (SimCLR Pre-train Model)</p>	

4.3 分類任務二：黃斑前膜與其他疾病之分類

在此任務中將使用「ERM vs Others diseases 資料集」來做模型的訓練與測試，共有 3748 張 OCT 影像，其中有 1874 張 OCT 影像來自 828 位患有 ERM 之病患；1874 張 OCT 影像來自 913 位黃斑部結構異常(不包含 ERM)之病患。其中 80% 用來做 training (2998 張)；10% 用來做 validation (375 張)；10% 用來做 testing (375 張)。在本任務中所使用的模型為 ImageNet Pre-train Model 及 SimCLR Pre-train Model，各項超參數設定如表 4.7 所示。

表 4.7 黃斑前膜與其他疾病之分類任務中各模型之超參數設定

	ImageNet Pre-train Model	SimCLR Pre-train Model
超參數	Epoch: 70 Batch size: 32 Optimizer: Adam Initial learning rate: 5×10^{-7} Weight decay: 1×10^{-4}	Epoch: 100 Batch size: 32 Optimizer: Adam Initial learning rate: 5×10^{-7} Weight decay: 1×10^{-4}

在表 4.8 中可以看到 ImageNet Pre-train Model 在 10 次交叉驗證訓練後其在測試集中平均準確率為 89.81%，以及平均 1.54% 的標準差；在表 4.9 中可以看到 SimCLR Pre-train Model 在 10 次交叉驗證訓練後其在測試集中平均準確率為 86.4%，以及平均 1.34% 的標準差。兩者皆能達到 85% 以上的準確率，皆能正確的辨別 ERM 與其他疾病的視網膜 OCT 影像，但同樣可以觀察到 ImageNet Pre-train Model 的表現依舊略優於 SimCLR Pre-train Model。

在表 4.10 中的 Grad-CAM 影像中可以觀察到 ImageNet Pre-train Model 與 SimCLR Pre-train Model 在看黃斑前膜影像時皆可以正確的將重點放到有 ERM 亮帶的部分，但 SimCLR Pre-train Model 除了該部分外還會將重點放在其他的部分；在表 4.11 中的 Grad-CAM 影像中可以觀察到 ImageNet Pre-train Model 與 SimCLR Pre-train Model 在看其他疾病影像時皆可以觀察到有異常的區域，但 SimCLR Pre-train Model 同樣會看除了異常區域外的其他部分。

表 4.8 ImageNet Pre-train Model 在黃斑前膜與其他疾病分類任務中測試集之表現結果

	1	2	3	4	5	6	7	8	9	10	Avg ±SD
Accuracy (%)	88.3	90.4	91.5	91.5	90.4	89.6	90.9	89.6	86.1	89.8	89.81 ±1.54

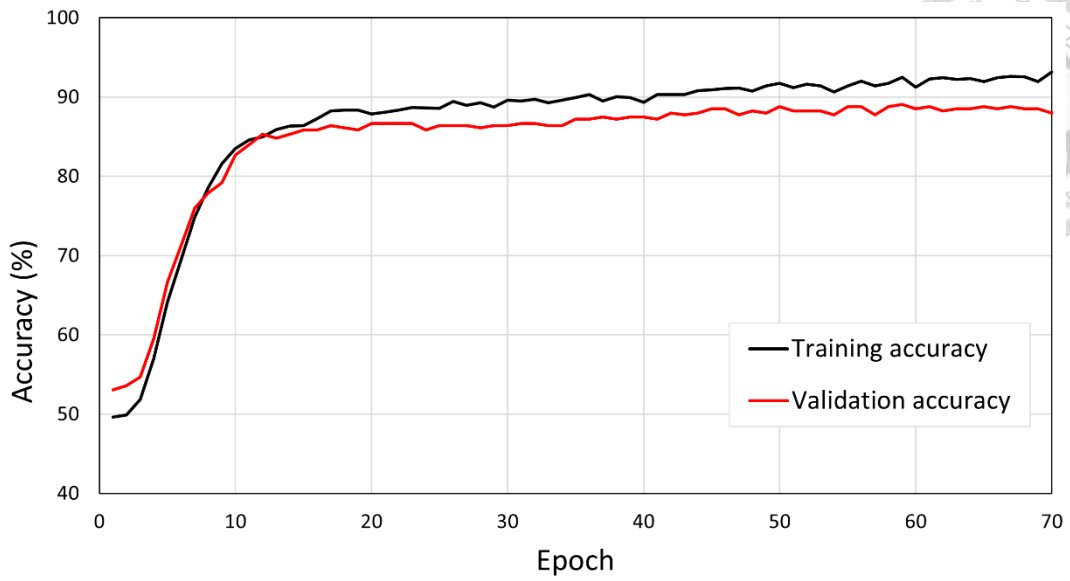


圖 4.8 ImageNet Pre-train Model 在黃斑前膜與其他疾病分類任務中訓練集與驗證集準確率與 Epoch 做圖結果

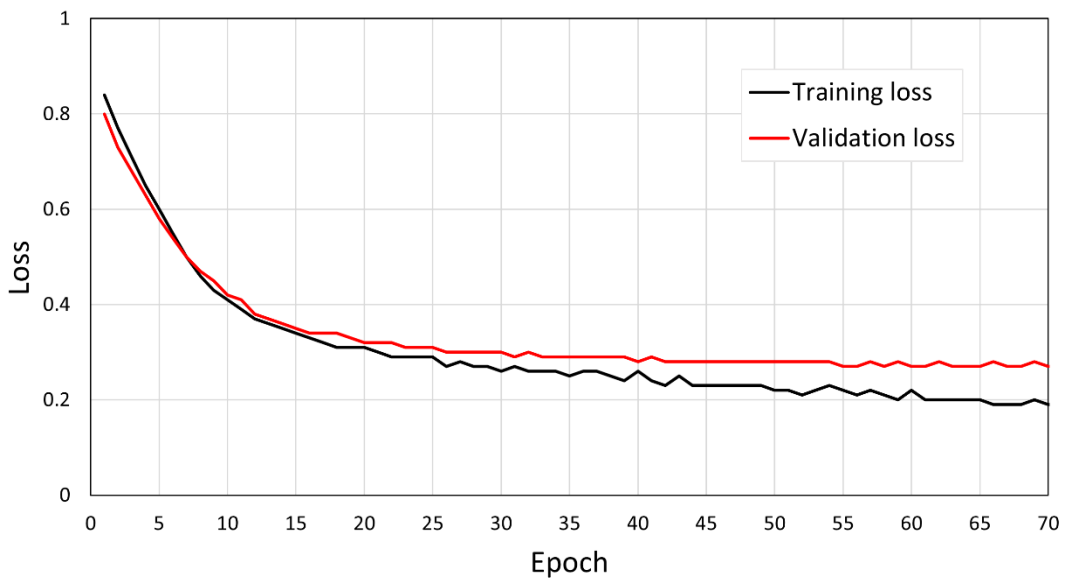


圖 4.9 ImageNet Pre-train Model 在黃斑前膜與其他疾病分類任務中訓練集與驗證集 Loss 與 Epoch 做圖結果

表 4.9 SimCLR Pre-train Model 在黃斑前膜與其他疾病分類任務中測試集之表現結果

	1	2	3	4	5	6	7	8	9	10	Avg ±SD
Accuracy (%)	84.8	87.7	87.7	87.5	86.1	86.1	88.5	86.1	84.2	85.3	86.4 ±1.34

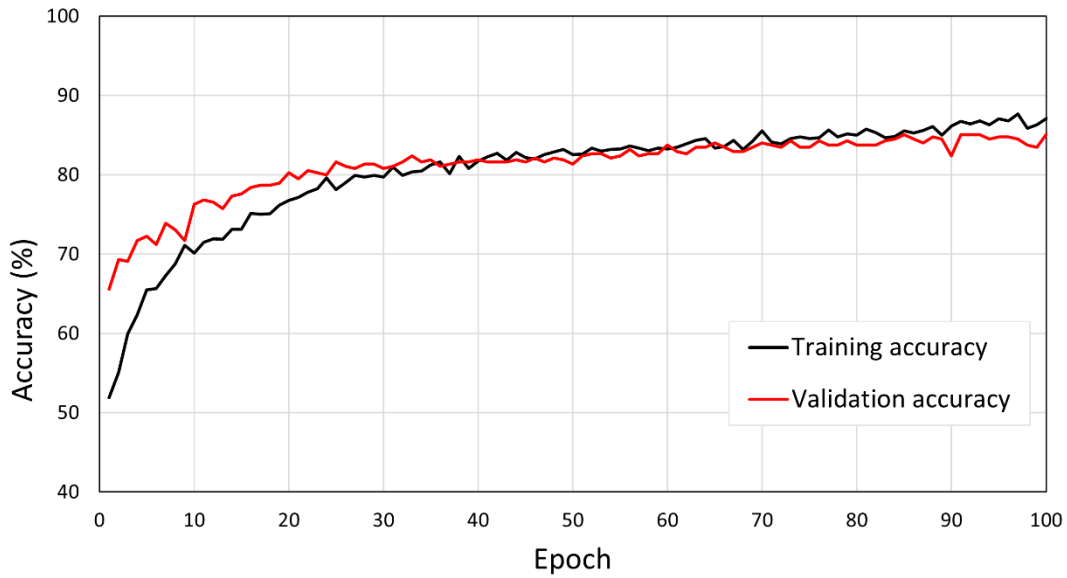


圖 4.10 SimCLR Pre-train Model 在黃斑前膜與其他疾病分類任務中訓練集與驗證集準確率與 Epoch 做圖結果

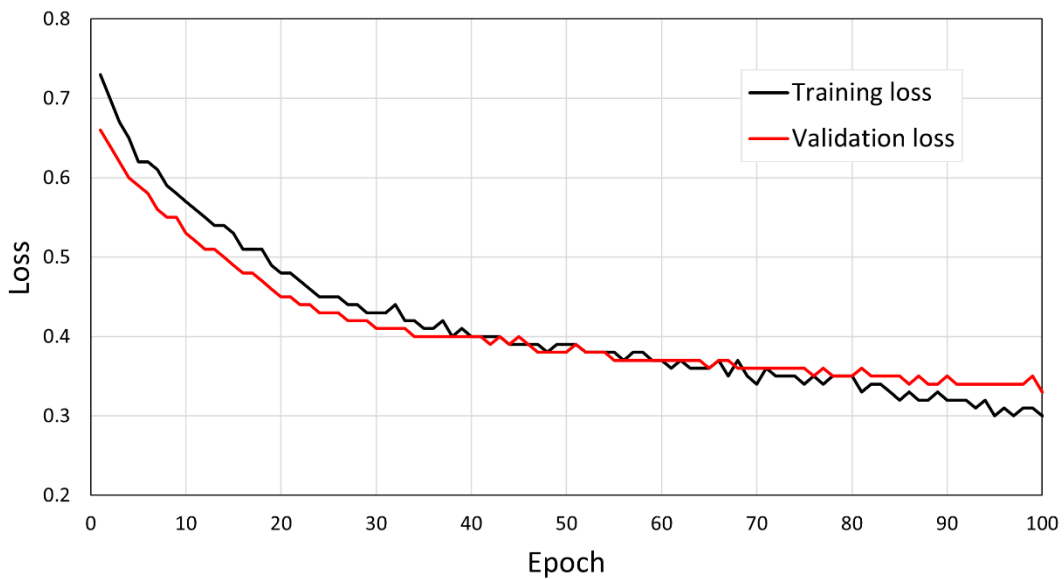


圖 4.11 SimCLR Pre-train Model 在黃斑前膜與其他疾病分類任務中訓練集與驗證集準確率與 Epoch 做圖結果

證集 Loss 與 Epoch 做圖結果

表 4.10 ImageNet Pre-train Model 及 SimCLR Pre-train Model 經黃斑前膜與其他疾病分類任務訓練後最後一層對黃斑前膜影像之 Grad-CAM 影像

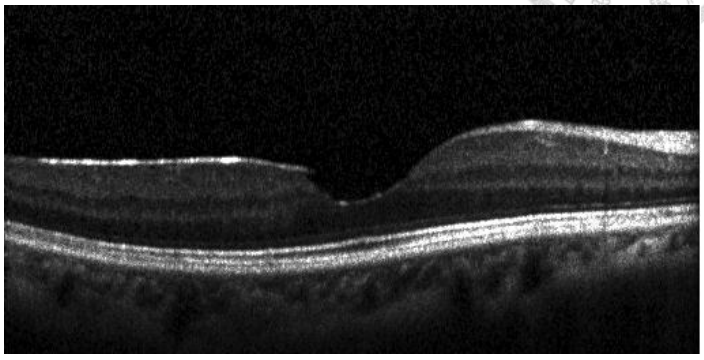
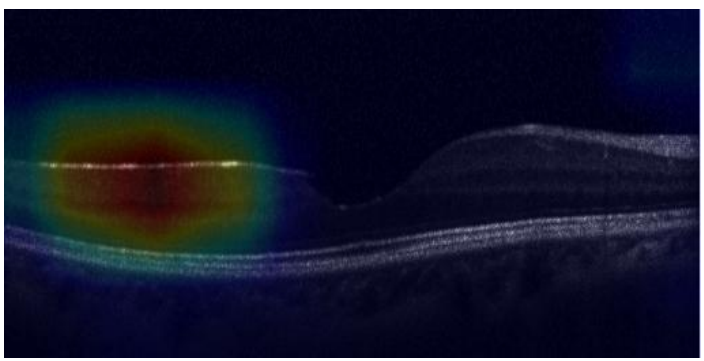
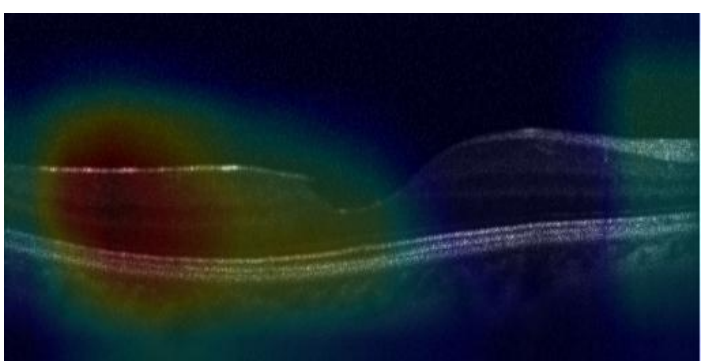
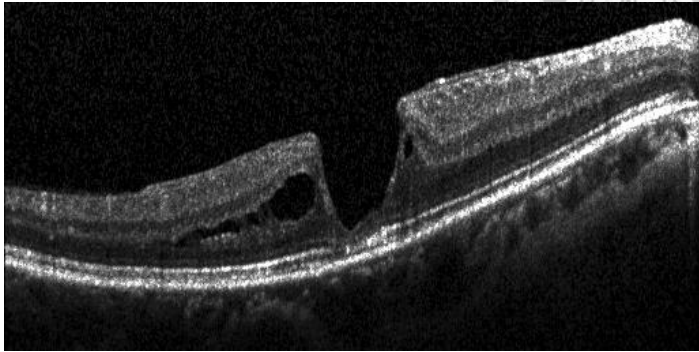
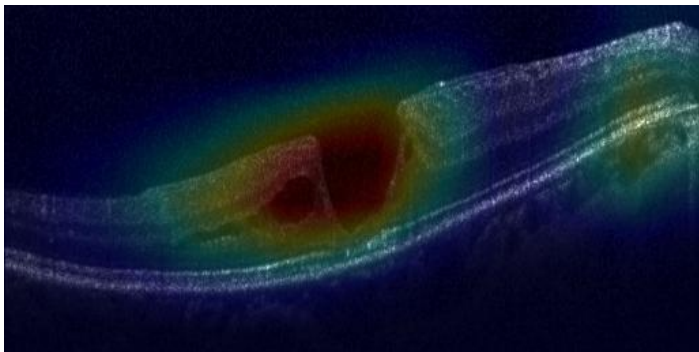
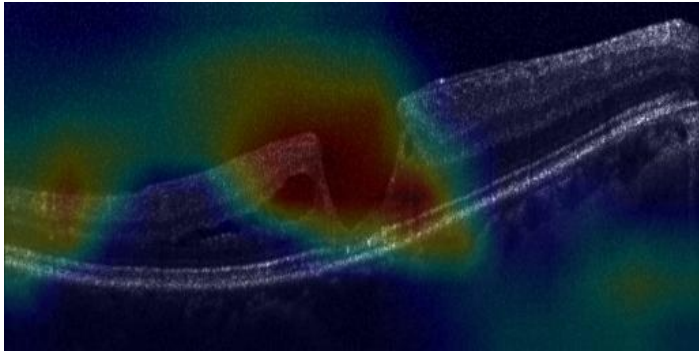
Input Image (ERM)	
Layer4 Grad-CAM (ImageNet Pre-train Model)	
Layer4 Grad-CAM (SimCLR Pre-train Model)	

表 4.11 ImageNet Pre-train Model 及 SimCLR Pre-train Model 經黃斑前膜與其他疾病分類任務訓練後最後一層對其他疾病影像之 Grad-CAM 影像

<p>Input Image (Other Diseases)</p>	
<p>Layer4 Grad-CAM (ImageNet Pre-train Model)</p>	
<p>Layer4 Grad-CAM (SimCLR Pre-train Model)</p>	

4.4 分類任務三：健康與黃斑前膜之分類

在此任務中將使用「ERM vs Healthy 資料集」來做模型的訓練與測試，共有 1142 張 OCT 影像，其中有 606 張 OCT 影像來自 556 位健康受試者；536 張 OCT 影像來自 254 位患有 ERM 之病患。其中 80% 用來做 training (914 張)；10% 用來做 validation (114 張)；10% 用來做 testing (114 張)。在本任務中所使用的模型為 ImageNet Pre-train Model 及 SimCLR Pre-train Model，各項超參數設定如表 4.12 所示。

表 4.12 健康與黃斑前膜之分類任務中各模型之超參數設定

	ImageNet Pre-train Model	SimCLR Pre-train Model
超參數	Epoch: 50 Batch size: 32 Optimizer: Adam Initial learning rate: 5×10^{-8} Weight decay: 1×10^{-4}	Epoch: 50 Batch size: 32 Optimizer: Adam Initial learning rate: 5×10^{-8} Weight decay: 1×10^{-4}

在表 4.13 中可以看到 ImageNet Pre-train Model 在 10 次交叉驗證訓練後其在測試集中平均準確率為 98.96%，以及平均 1.34% 的標準差；在表 4.14 中可以看到 SimCLR Pre-train Model 在 10 次交叉驗證訓練後其在測試集中平均準確率為 99.38%，以及平均 1.04% 的標準差。兩者之準確率皆可以到達約 100%，皆能正確的辨別 ERM 與健康的視網膜 OCT 影像，在此分類任務中可以觀察到 SimCLR Pre-train Model 的表現首次略優於 ImageNet Pre-train Model。

在表 4.15 中的 Grad-CAM 影像中可以觀察到 ImageNet Pre-train Model 與 SimCLR Pre-train Model 在看健康影像時皆可以正確的將重點放到中央凹的部分，但 SimCLR Pre-train Model 除了該部分外還會將重點放在兩側；在表 4.16 中的 Grad-CAM 影像中可以觀察到 ImageNet Pre-train Model 與 SimCLR Pre-train Model 在看黃斑前膜影像時皆可以觀察到有有黃斑前膜的亮帶區域，但 SimCLR Pre-train Model 同樣會看除了該區域外的其他部分。

表 4.13 ImageNet Pre-train Model 在健康與黃斑前膜分類任務中測試集之表現結果

	1	2	3	4	5	6	7	8	9	10	Avg ±SD
Accuracy (%)	100	100	96.5	97.4	97.4	100	100	98.3	100	100	98.96 ±1.34

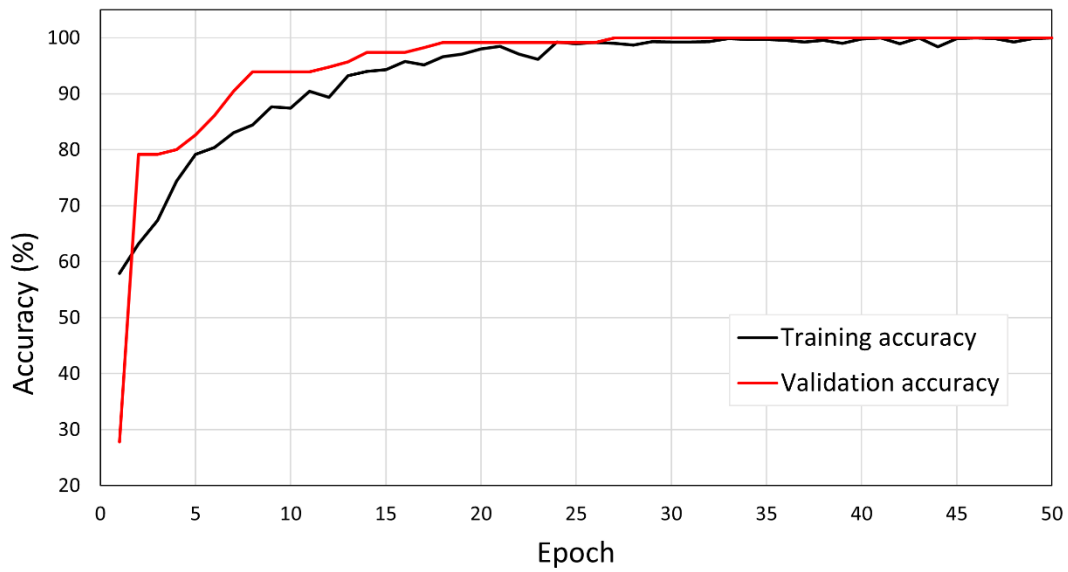


圖 4.12 ImageNet Pre-train Model 在健康與黃斑前膜分類任務中訓練集與驗證集準確率與 Epoch 做圖結果

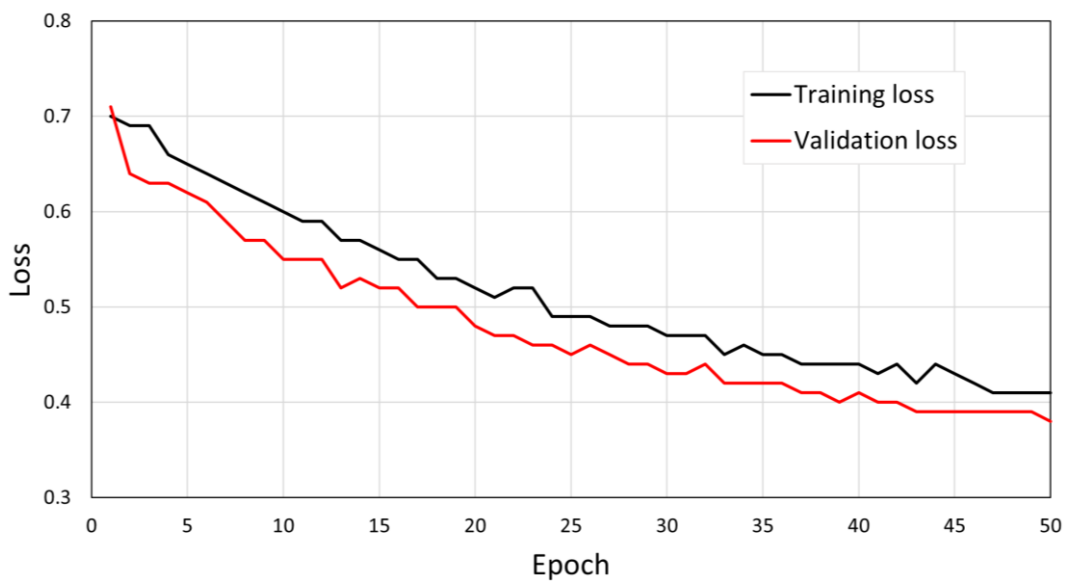


圖 4.13 ImageNet Pre-train Model 在健康與黃斑前膜分類任務中訓練集與驗證集 Loss 與 Epoch 做圖結果

表 4.14 SimCLR Pre-train Model 在健康與黃斑前膜分類任務中測試集之表現結果

	1	2	3	4	5	6	7	8	9	10	Avg ±SD
Accuracy (%)	100	99.1	100	96.5	100	100	100	100	99.1	99.1	99.38 ±1.04

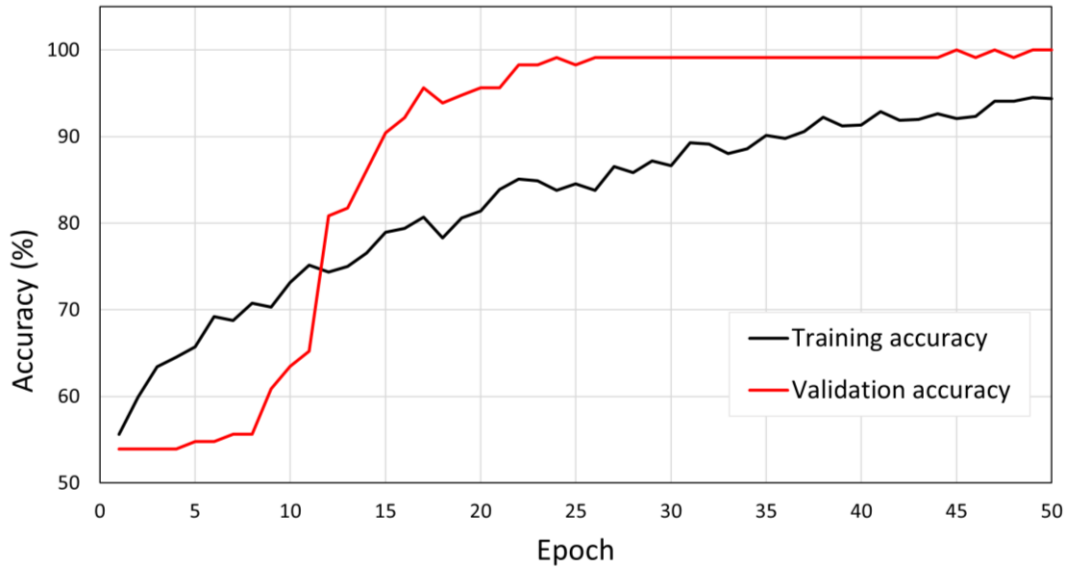


圖 4.14 SimCLR Pre-train Model 在健康與黃斑前膜分類任務中訓練集與驗證集準確率與 Epoch 做圖結果

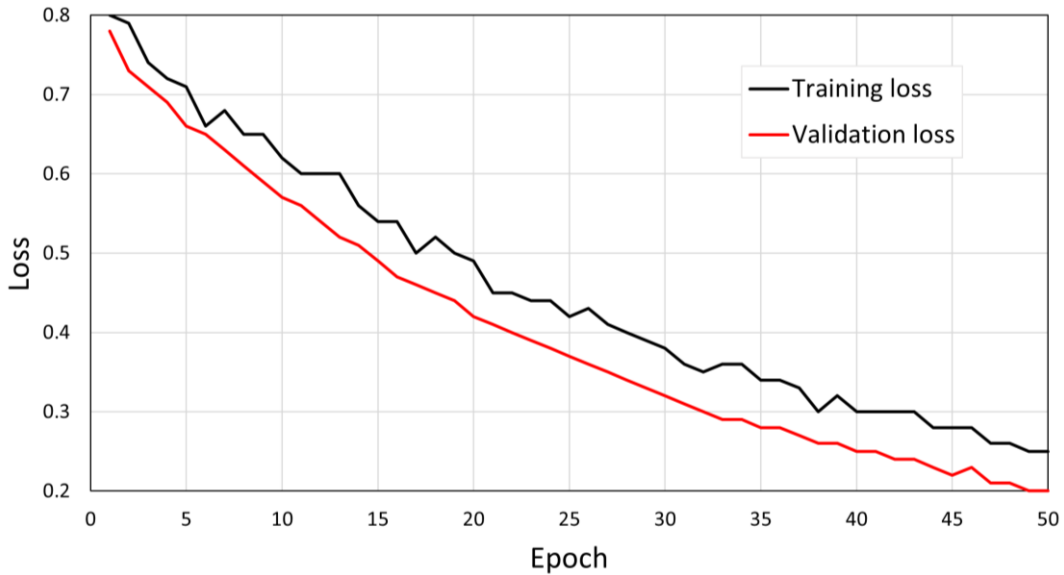


圖 4.15 SimCLR Pre-train Model 在健康與黃斑前膜分類任務中訓練集與驗證集 Loss 與 Epoch 做圖結果

表 4.15 ImageNet Pre-train Model 及 SimCLR Pre-train Model 經健康與黃斑前膜分類任務訓練後最後一層對健康影像之 Grad-CAM 影像

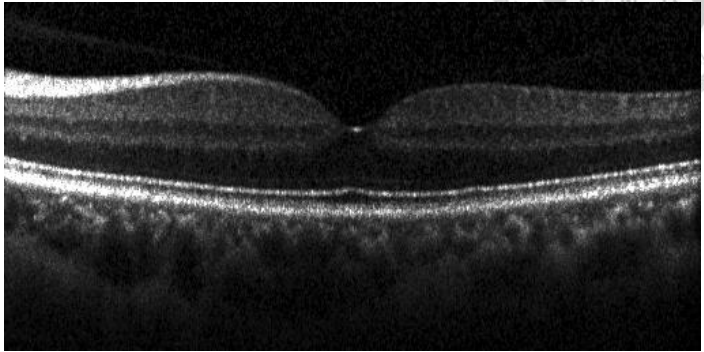
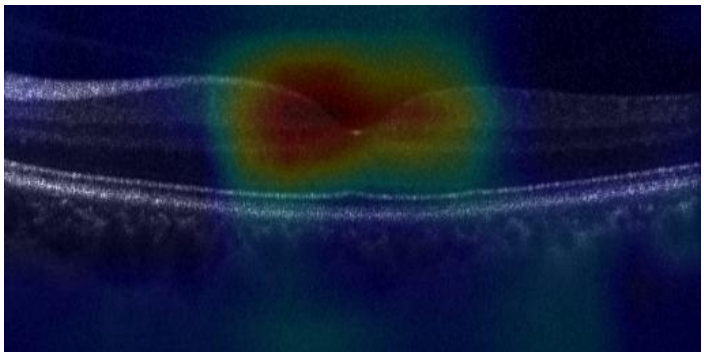
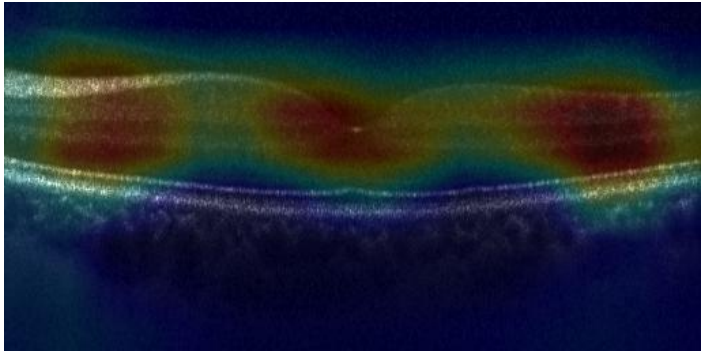
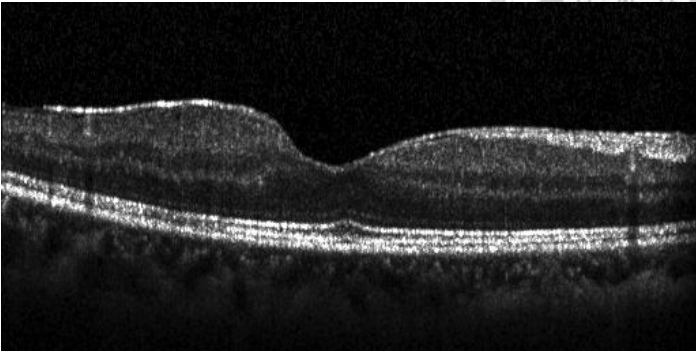
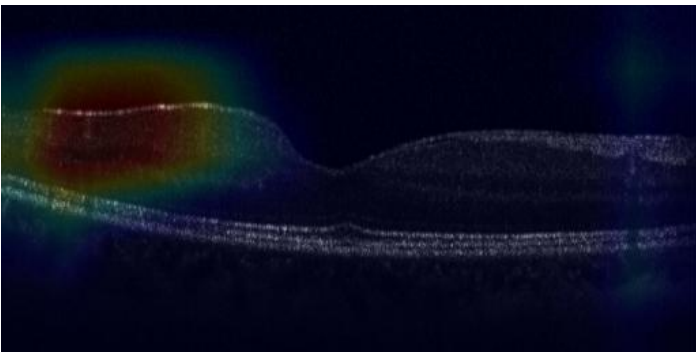
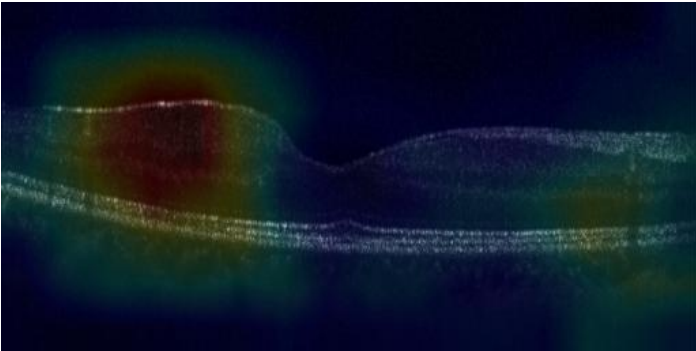
<p>Input Image (Healthy)</p>	
<p>Layer4 Grad-CAM (ImageNet Pre-train Model)</p>	
<p>Layer4 Grad-CAM (SimCLR Pre-train Model)</p>	

表 4.16 ImageNet Pre-train Model 及 SimCLR Pre-train Model 經健康與黃斑前膜分類任務訓練後最後一層對黃斑前膜影像之 Grad-CAM 影像

<p>Input Image (ERM)</p>	
<p>Layer4 Grad-CAM (ImageNet Pre-train Model)</p>	
<p>Layer4 Grad-CAM (SimCLR Pre-train Model)</p>	

4.5 分類任務四：健康與亞健康之分類

在此任務中將使用「Healthy vs Sub-healthy(Female)資料集」來做模型的訓練與測試，共有 764 張 OCT 影像，其中有 400 張 OCT 影像來自 303 位健康受試者；364 張 OCT 影像來自 162 位患有 ERM 之病患，在此資料集中只使用女性影像是與謝易庭醫師討論後認為不單男女之間視網膜影像本身就存在差異，且黃斑前膜相對好發於女性，因此醫師建議可以將男女分開做訓練[33]。其中 80%用來做 training (611 張)；10%用來做 validation (76 張)；10%用來做 testing (77 張)。在本任務中所使用的模型為 ImageNet Pre-train Model、SimCLR Pre-train Model 及

OCT Pre-train Model，各項超參數設定如表 4.17 所示。

表 4.17 健康與黃斑前膜之分類任務中各模型之超參數設定

	ImageNet Pre-train Model	SimCLR Pre-train Model	OCT Pre-train Model
超參數	Epoch: 150 Batch size: 64 Optimizer: Adam Initial learning rate: 7×10^{-6} Weight decay: 1×10^{-4}	Epoch: 150 Batch size: 64 Optimizer: Adam Initial learning rate: 7×10^{-6} Weight decay: 1×10^{-4}	Epoch: 150 Batch size: 64 Optimizer: Adam Initial learning rate: 7×10^{-6} Weight decay: 1×10^{-4}

在表 4.18 中可以看到 ImageNet Pre-train Model 在 10 次交叉驗證訓練後其在測試集中平均準確率為 65.99%，以及平均 7.93% 的標準差；在表 4.19 中可以看到 SimCLR Pre-train Model 在 10 次交叉驗證訓練後其在測試集中平均準確率為 65.92%，以及平均 9.9% 的標準差；在表 4.20 中可以看到 OCT Pre-train Model 在 10 次交叉驗證訓練後其在測試集中平均準確率為 66.6%，以及平均 6.09% 的標準差。整題而言，三種模型的平均準確率都約在 66% 左右，雖然相比隨機猜測的 50% 優秀，但相較於前三類任務，在本分類任務中仍無法明確區分兩類影像，顯示分類難度較高。此外，從標準差的結果可看出三種模型的表現波動皆較大，表示其穩定性仍有限。進一步比較三種模型的結果，OCT Pre-train Model 擁有最高的平均準確率及最低的平均標準差，顯示其在準確度與穩定性方面皆優於其他模型。

表 4.18 ImageNet Pre-train Model 在健康與亞健康分類任務中測試集之表現結果

	1	2	3	4	5	6	7	8	9	10	Avg ±SD
Accuracy (%)	66.7	53.1	78.4	71.4	57	74.5	56.2	65.1	71.4	66.1	65.99 ±7.93

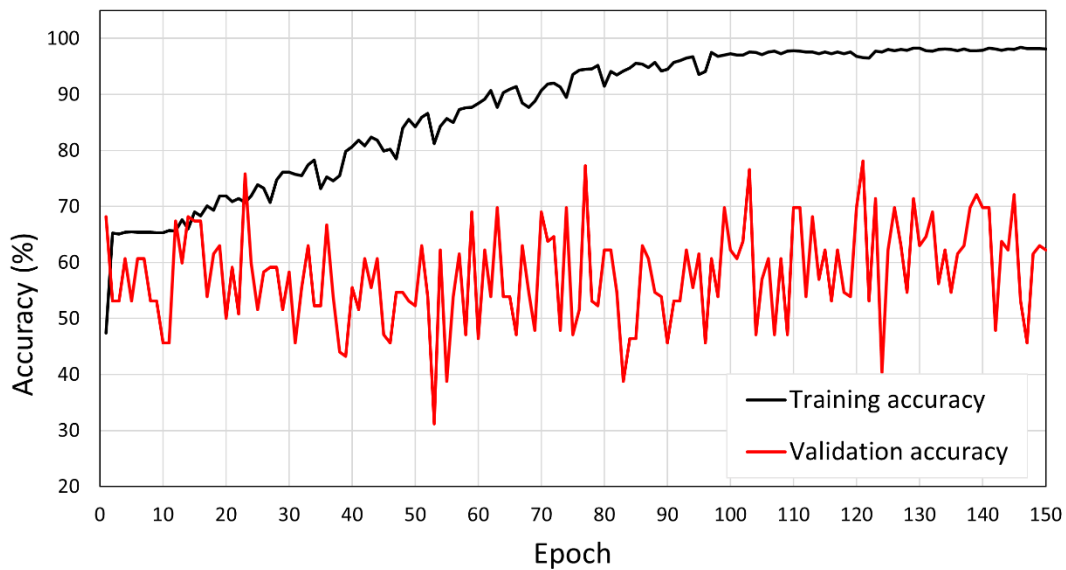


圖 4.16 ImageNet Pre-train Model 在健康與亞健康分類任務中訓練集與驗證集準確率與 Epoch 做圖結果

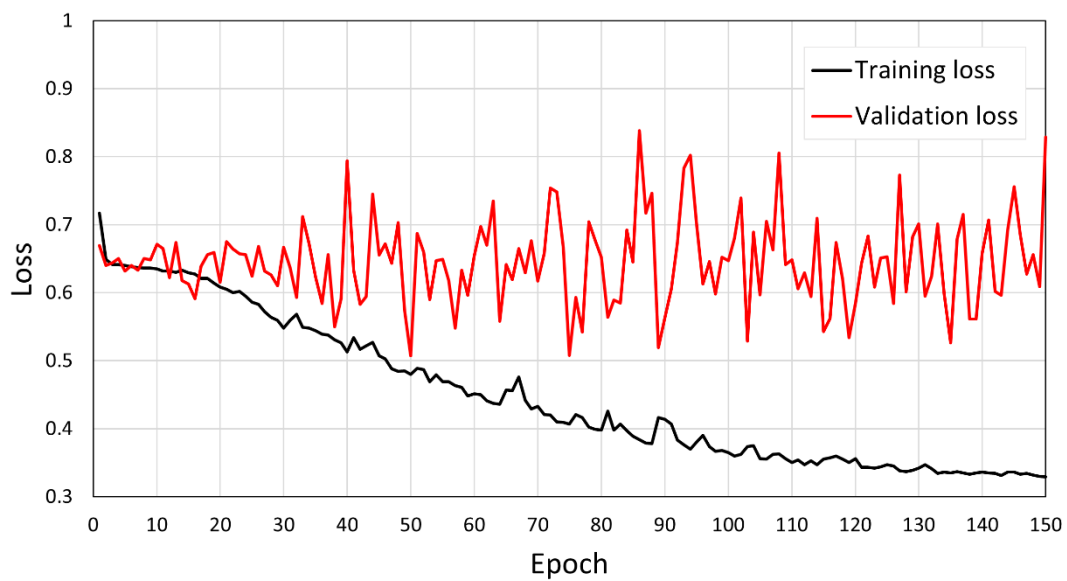


圖 4.17 ImageNet Pre-train Model 在健康與亞健康分類任務中訓練集與驗證集 Loss 與 Epoch 做圖結果

表 4.19 SimCLR Pre-train Model 在健康與亞健康分類任務中測試集之表現結果

	1	2	3	4	5	6	7	8	9	10	Avg ±SD
Accuracy (%)	57.6	50	78.4	72.1	73.7	70.6	48.7	69	64.6	74.5	65.92 ±9.9

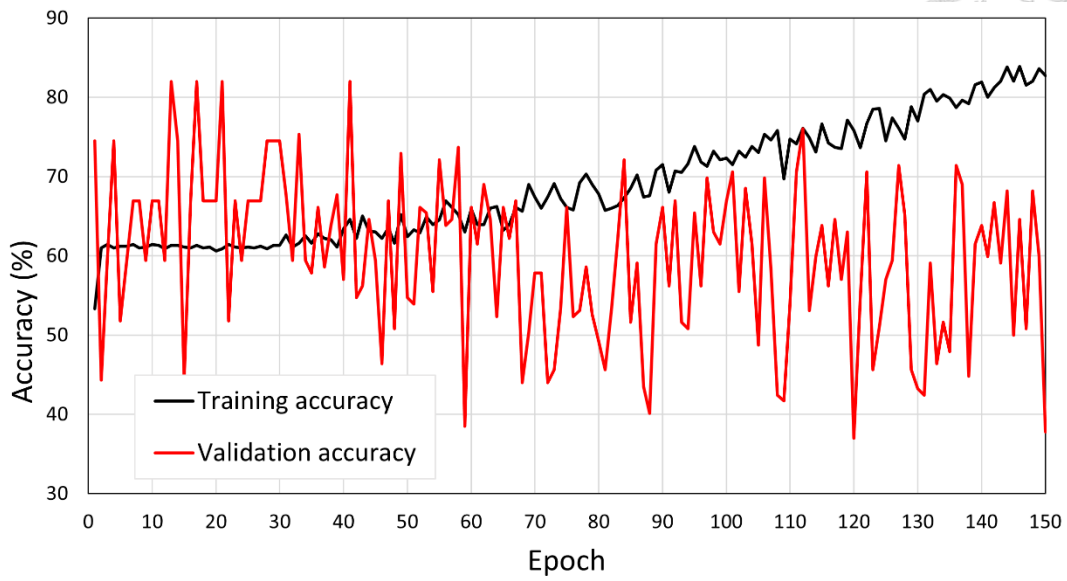


圖 4.18 SimCLR Pre-train Model 在健康與亞健康分類任務中訓練集與驗證集準確率與 Epoch 做圖結果

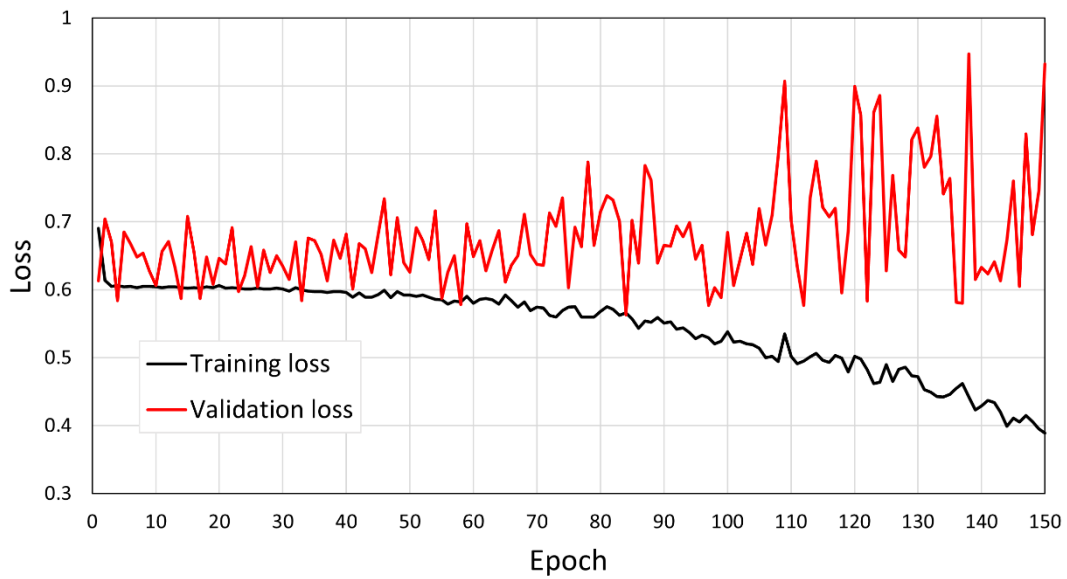


圖 4.19 SimCLR Pre-train Model 在健康與亞健康分類任務中訓練集與驗證集 Loss 與 Epoch 做圖結果

表 4.20 OCT Pre-train Model 在健康與亞健康分類任務中測試集之表現結果[33]

	1	2	3	4	5	6	7	8	9	10	Avg ±SD
Accuracy (%)	62.8	51.7	74.1	67.9	71.0	68.8	72.9	66.7	63.2	68.9	66.60 ±6.09

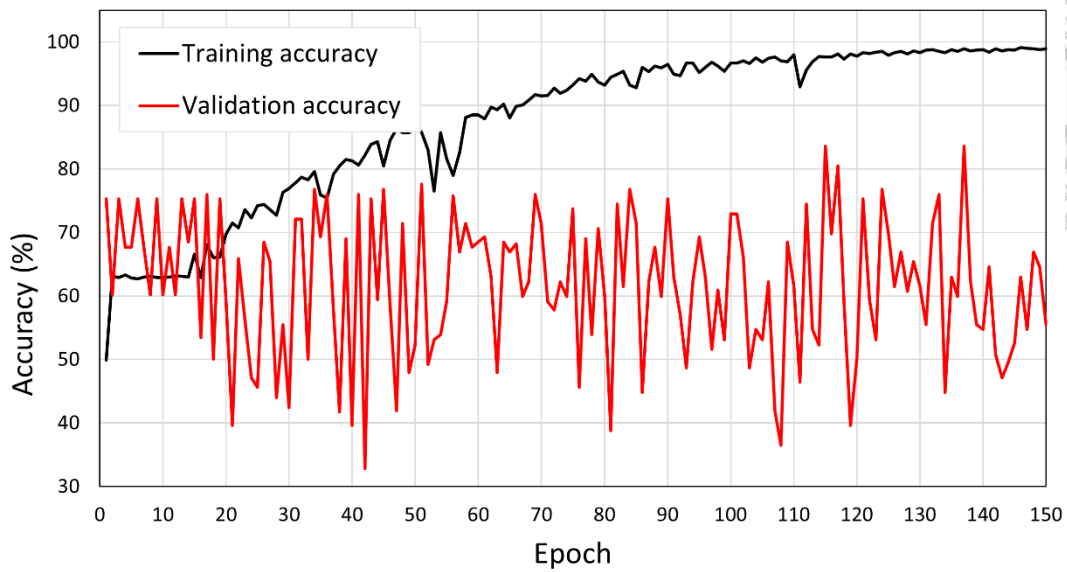


圖 4.20 OCT Pre-train Model 在健康與亞健康分類任務中訓練集與驗證集準確率與 Epoch 做圖結果[33]

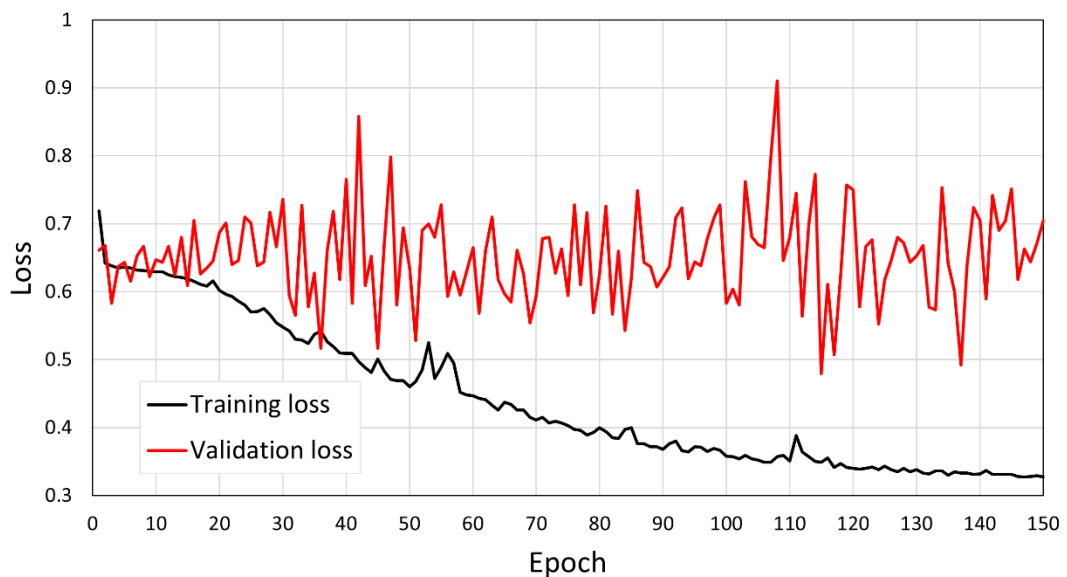


圖 4.21 OCT Pre-train Model 在健康與亞健康分類任務中訓練集與驗證集 Loss 與 Epoch 做圖結果[33]

4.6 各項分類任務之分析與討論

從以上的訓練結果中可以觀察到，SimCLR Pre-train Model 的表現僅在「健康與黃斑前膜」的分類任務中出現些微優勢，在其他任務中則皆不如 ImageNet Pre-train Model。

推測雖然 SimCLR 預訓練階段的 Loss 降低明顯，且驗證集中準確率似乎良好，但這未必代表模型學到了與臨床分類任務相關的語意特徵(Semantic Features，例如視網膜層間的變形或病灶區的邊界變化)。由於 SimCLR 是透過資料增強(Augmentation)來產生正負樣本對，模型往往傾向於學習最容易降低 Loss 的方式，例如依靠亮度、對比度、紋理等非結構性差異進行判斷，而非真正學習病灶結構。這種現象與 Geirhos 等人所提出的「Shortcut Learning」概念一致，即神經網路傾向依賴輸入中的簡單但不具泛化性的特徵進行預測，而忽略任務真正相關的語意結構[34]。

雖然本研究使用了約兩萬張 OCT 影像來進行 SimCLR 的自監督預訓練，但仍屬於中小型資料集規模。Azizi 等人指出，自監督模型僅當預訓練資料量達到數十萬甚至百萬張時，才能穩定學得泛化性強的表徵，進而在後續任務中超越監督式預訓模型；若資料不足，則表現容易不穩定甚至低於傳統預訓模型[35]。

因此，本研究中僅使用約兩萬張 OCT 影像進行 SimCLR 預訓練，可能尚不足以讓模型學到影像真正的語意結構，導致後續任務表現不佳。反觀 ImageNet Pre-train Model，雖然其來源影像並非 OCT 影像，但由於其資料量龐大(約 130 萬張)、涵蓋影像多樣性高，使模型得以學習到穩定的低階結構特徵，例如邊界偵測與紋理抽取。由於視網膜病變(如 ERM)常呈現在層與層之間的邊界異常，因此這類特徵在 OCT 分類任務中仍具有良好泛化能力，並在本實驗中展現出更穩定的分類表現。

除此之外，OCT Pre-train Model 雖然後續是微調在男女的分類任務上，但在本實驗中仍然在「健康與亞健康」中有著最優異的表現，推測在中小型資料集規模下雖然使用自監督學習的效果有限，但監督式微調即使所進行的任務與最終任務不同，但使用相同格式的影像資料，仍然可以幫助模型更好的學習到影像中的語意結構，進而提升在其他任務中的表現。

在前三項任務的 Grad-CAM 影像中可以觀察到，ImageNet Pre-train Model 及 SimCLR Pre-train Model 不論是在健康影像、黃斑前膜影像、其他疾病影像中皆可以成功聚焦於各類影像的關鍵區域，但 SimCLR Pre-train Model 不管在何種任務中皆會額外關注非該類別特徵的部分。推測與前面提到之「Shortcut Learning」有關，由於在本實驗中用來做 SimCLR 自監督學習之資料集規模有限，導致模型在預訓期間傾向學習較易辨識但與語意無關的特徵作為捷徑，雖然透過後續的監督式微調能引導模型學習正確的語意區域，但其注意力仍可能受這些捷徑特徵影響。因此即使 SimCLR Pre-train Model 型在測試集上仍具備分類能力，其整體表現仍略遜於 ImageNet Pre-train Model。

Chapter 5 聚類分析與潛在亞健康樣本之探討

在第四章中進行了健康與亞健康之分類任務，雖採用了三種不同的預訓練方式，但最終結準確率卻都約為 66%，認為可能的原因出在健康影像的標註上。在本實驗中對於黃斑前膜亞健康影像的標註需要另一隻眼睛被確診為黃斑前膜才行，但實際上想要分辨出健康影像與亞健康影像的目的是為了找出較容易發病之眼睛已達到預防的效果，而有些病患可能兩隻眼睛皆還為健康眼睛還未發病，但其中一隻或兩隻眼睛皆為「亞健康」狀態了。故在本章中將會使用無監督聚類分析與半監督聚類分析的方式來找出健康資料集中潛在的亞健康樣本。

5.1 聚類分析設計與方法

本實驗使用聚類(Clustering)分析這類機器學習方式，主要是因為與深度學習中的監督式分類相比較不依賴人工標註。兩種方式雖然皆可以達到分類影像的目的，但深度學習訓練模型的過程中完全依賴影像的標註，不論影像標註是否正確皆會強迫模型學習，導致過擬合或學不到具有代表性的特徵。反之聚類分析則可以根據影像本身的結構特徵來進行分群不依賴人工標註，且像本實驗中只對亞健康影像之標註較有信心時也可以使用半監督聚類分析，半聚類分析可以在只提供一類標註的情況下引導模型學習該類之語意特徵，不像深度學習的監督式分類想要讓模型學會幾類別之分類就需要提供多少類別之標註。故最終決定使用聚類分析的方式來找出健康資料集中潛在的亞健康樣本。

在無監督聚類分析中將使用 K-means 演算法，而在半監督聚類分析終將使用 Semi-supervised Deep Embedded Clustering (SDEC)演算法。在兩種不同的聚類分析中皆使用「Healthy vs Sub-healthy(Female)資料集」並使用 ImageNet Pre-train Model、SimCLR Pre-train Model、OCT Pre-train Model 這三種模型重複進行十次，並將健康影像被各種模型認為是亞健康樣本的次數紀錄，以此方法初步從健康資料集中挑選出潛在的亞健康樣本再進行後續之分析。除此之外，在這十次聚類分析中會記錄混淆矩陣(Confusion Matrix)分別計算出準確率(Accuracy)、精確率(Precision)、召回率(Recall)、F1 分數(F1 Score)、特異度(Specificity)、假陽性率(False Positive Rate; FPR)如式(5.1)一式(5.6)所示[36]，在實驗中以上數據皆使用亞健康做為正類(Positive)計算。

	實際 sub-healthy	實際 healthy
預測 sub-healthy	TP	FP
預測 healthy	FN	TN

圖 5.1 混淆矩陣示意圖[36]

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \quad (5.1)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (5.2)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (5.3)$$

$$\text{F1 Score} = \frac{2}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}} \quad (5.4)$$

$$\text{Specificity} = \frac{TN}{FP+TN} \quad (5.5)$$

$$\text{FPR} = \frac{FP}{FP+TN} \quad (5.6)$$

5.2 無監督聚類分析

圖 5.2—圖 5.4 可以看到在三個不同模型左側預測分群中將資料明確的分為了兩群，但在右側真實標註分布中卻沒有像預測分群一樣呈現清晰的兩個類別，也沒有部分健康樣本向亞健康樣本聚集的趨勢，無法幫助選出健康資料集中潛在的亞健康樣本。為了進一步挑選潛在亞健康樣本，後續將採取不同的策略進行。將此無監督聚類分析執行十次，紀錄標註為健康樣本但被模型預測為亞健康的樣本，透過統計這些樣本在多次分群中被誤歸類的次數，判斷其是否為潛在的亞健康樣本。

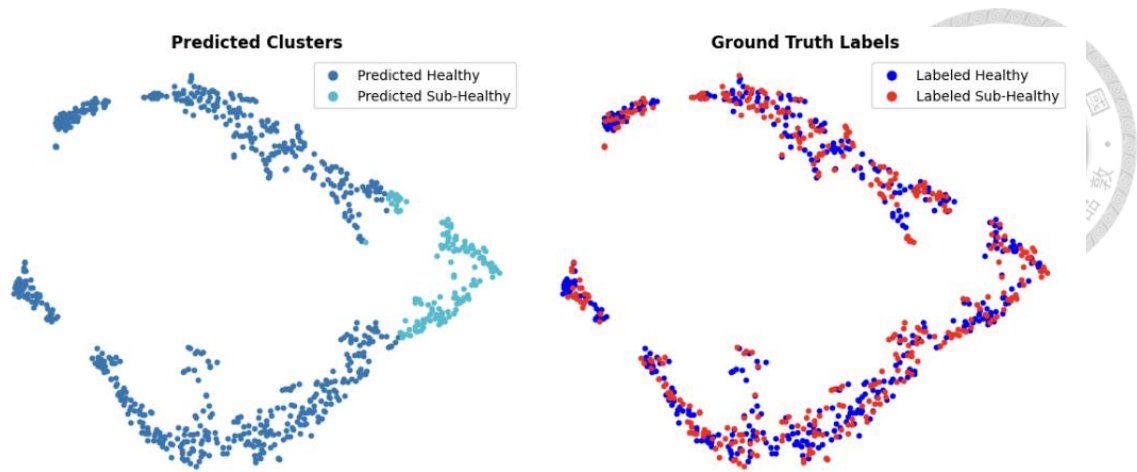


圖 5.2 ImageNet Pre-train Model 無監督聚類分析預測分群與實際分群結果

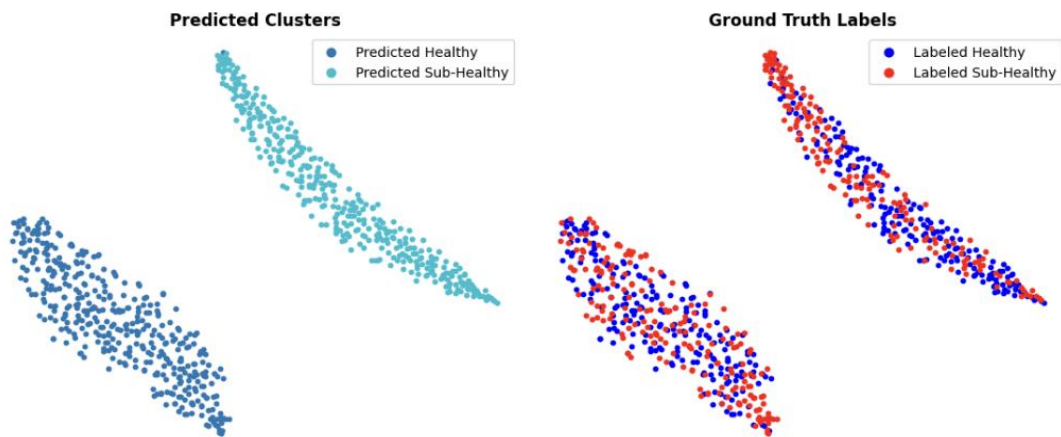


圖 5.3 SimCLR Pre-train Model 無監督聚類分析預測分群與實際分群結果

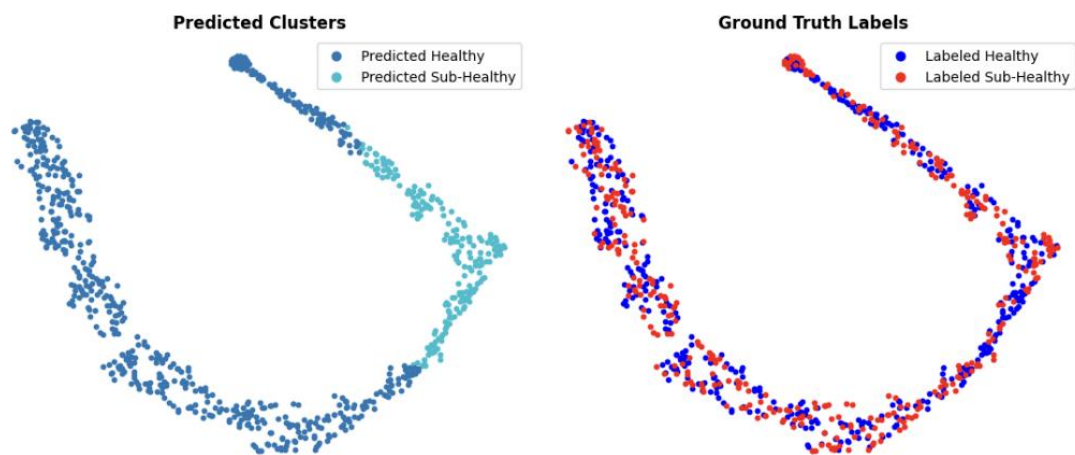


圖 5.4 OCT Pre-train Model 無監督聚類分析預測分群與實際分群結果

在表 5.1 中可以觀察到，ImageNet Pre-train Model 的 FPR 較低，顯示其在判斷亞健康樣本時相對謹慎。然而在表 5.2 中進一步分析被重複挑選為潛在亞健康樣本的次數後發現該模型中樣本僅最多被挑選 6 次，顯示儘管其分類較為保

守，但最終挑選出來的潛在亞健康樣本具有較低的穩定性與一致性；相較之下，SimCLR Pre-train Model 與 OCT Pre-train Model 對於亞健康樣本的判斷則明顯較為積極。在每次訓練中，這兩種模型都在健康資料集中預測大量的亞健康樣本。在統計分析中，SimCLR Pre-train Model 中有 1 個樣本在 10 次分群中都被選中、44 個樣本出現 9 次，而 OCT Pre-train Model 則有 46 個樣本出現 8 次。雖然這兩種模型在潛在亞健康樣本的挑選上展現出較高的重複性與一致性，但由於每次選出的樣本數量過多，使得高重複次數的樣本數量也偏高。綜合以上觀察結果認為僅依賴無監督式聚類分析，難以有效且穩定地辨識健康資料集中潛在的亞健康樣本，故接下來將使用半聚類分析的方式來挑選出潛在的亞健康樣本。

表 5.1 ImageNet Pre-train Model 在 Healthy vs Sub-healthy(Female)資料集無監督聚類分析結果

	1	2	3	4	5	6	7	8	9	10	Avg ±SD
Accuracy(%)	53	52.1	53.1	52.1	50.9	52.2	53	52	49.9	52.1	52.04 ±0.94
Precision(%)	52.6	49.2	52.1	46.4	46.1	49.7	52.9	49	41.4	49.3	48.87 ±3.33
Recall(%)	13.7	17.9	20.9	3.6	17.9	20.1	12.6	20.3	12.6	20.1	15.97 ±5.15
F1 score(%)	21.7	26.2	29.8	6.7	25.8	28.6	20.4	28.7	19.3	28.6	23.58 ±6.66
Specificity(%)	88.75	83.25	82.5	96.25	81	81.5	89.75	80.75	83.75	81.25	84.875 ±4.84
FPR(%)	11.25	16.75	17.5	3.75	19	18.5	10.25	19.25	16.25	18.75	15.125 ±4.84

表 5.2 ImageNet Pre-train Model 在無監督聚類分析選出潛在亞健康樣本統計結果

Filename	Count
0_00086033_11.png	6
0_00087232_11.png	6
0_00085015_11.png	6
0_00021034_11.png	6

表 5.3 SimCLR Pre-train Model 在 Healthy vs Sub-healthy(Female)資料集無監督聚類分析結果

	1	2	3	4	5	6	7	8	9	10	Avg ±SD
Accuracy(%)	50.4	47.6	50.8	50.1	50	50.3	51	49.6	50.4	54.1	50.43 ±1.51
Precision(%)	47.9	45.2	47.8	47.4	47.5	47.6	48.6	47.2	47.5	53.1	47.98 ±1.9
Recall(%)	47.8	46.2	36.3	42	47	44.5	47.5	48.6	39	31	42.99 ±5.56
F1 score(%)	47.8	45.7	41.3	44.5	47.2	46	48	47.9	42.8	39.1	45.03 ±2.92
Specificity(%)	52.75	49	64	57.5	52.75	55.5	54.25	50.5	60.75	75	57.2 ±7.34
FPR(%)	47.25	51	36	42.5	47.25	44.5	45.75	49.5	39.25	25	42.8 ±7.34

表 5.4 SimCLR Pre-train Model 在無監督聚類分析選出潛在亞健康樣本統計結果

Filename	Count	Filename	Count
0_00103124_11.png	10	0_00101205_11.png	9
0_00114324_11.png	9	0_00039343_11.png	9
0_00117181_11.png	9	0_00054220_11.png	9
0_00104960_11.png	9	0_00066454_11.png	9
0_00112325_11.png	9	0_00050172_11.png	9
0_00108438_11.png	9	0_00061592_11.png	9
0_00106446_11.png	9	0_00065338_11.png	9
0_00061144_11.png	9	0_00017916_11.png	9
0_00065948_11.png	9	0_00015147_11.png	9
0_00025180_11.png	9	0_00082499_11.png	9
0_00022763_11.png	9	0_00092845_11.png	9
0_00011540_11.png	9	0_00020741_11.png	9
0_00004444_11.png	9	0_00027737_11.png	9
0_00072856_11.png	9	0_00040698_11.png	9
0_00084155_11.png	9	0_00043556_11.png	9
0_00078868_11.png	9	0_00088041_11.png	9

0_00079688_11.png	9	0_00088675_11.png	9
0_00081089_11.png	9	0_00009731_11.png	9

表 5.5 OCT Pre-train Model 在 Healthy vs Sub-healthy(Female)資料集無監督聚類分析結果

	1	2	3	4	5	6	7	8	9	10	Avg ±SD
Accuracy(%)	51.2	44.8	57.6	49.5	51.2	51.4	49.9	50.3	43.5	50.3	49.97 ±3.64
Precision(%)	47.7	48.2	61.7	46.6	48.1	48.7	47	47.5	46.8	47.3	48.96 ±4.29
Recall(%)	25.3	38.4	55.6	41.5	31.6	35.2	40.9	42.3	38.1	38.7	38.76 ±7.44
F1 score(%)	33.1	42.7	58.5	43.9	38.1	40.9	43.7	44.7	42	42.6	43.02 ±6.1
Specificity(%)	74.75	52.25	60	56.75	69	66.25	58	57.5	49.75	60.75	60.5 ±7.22
FPR(%)	25.25	47.75	40	43.25	31	33.75	42	42.5	50.25	39.25	39.5 ±7.22

表 5.6 OCT Pre-train Model 在無監督聚類分析選出潛在亞健康樣本統計結果

Filename	Count	Filename	Count
0_00068590_11.png	8	0_00112416_11.png	8
0_00053897_11.png	8	0_00111348_11.png	8
0_00073292_11.png	8	0_00106859_11.png	8
0_00071620_11.png	8	0_00110827_11.png	8
0_00077152_11.png	8	0_00108475_11.png	8
0_00082314_11.png	8	0_00084155_11.png	8
0_00080541_11.png	8	0_00092845_11.png	8
0_00085486_11.png	8	0_00087204_11.png	8
0_00083362_11.png	8	0_00086852_11.png	8
0_00116277_11.png	8	0_00086575_11.png	8
0_00107473_11.png	8	0_00098270_11.png	8
0_00098634_11.png	8	0_00061895_11.png	8
0_00091303_11.png	8	0_00061442_11.png	8

0_00087466_11.png	8	0_00061363_11.png	8
0_00053174_11.png	8	0_00059001_11.png	8
0_00094456_11.png	8	0_00000613_11.png	8
0_00094914_11.png	8	0_00004444_11.png	8
0_00097022_11.png	8	0_00026477_11.png	8
0_00046519_11.png	8	0_00031071_11.png	8
0_00040698_11.png	8	0_00017916_11.png	8
0_00065338_11.png	8	0_00028617_11.png	8
0_00039063_11.png	8	0_00026872_11.png	8
0_00098634_11.png	8	0_00061895_11.png	8

5.3 半監督聚類分析

在半監督聚類分析的各项超參數設定如表 5.5 所示，ImageNet Pre-train Model、SimCLR Pre-train Model、OCT Pre-train Model 皆使用相同的超參數。半監督聚類分析在進行分群時仍然會微調模型的權重，以此來讓模型能夠獲得更適合分群的特徵向量，因此與無監督聚類分析不同需要額外設定超參數，而在半監督聚類分析中則需要。且在本實驗中只給予亞健康影像的標註資訊，希望能夠藉此引導模型學會亞健康影像之特徵，並最終能夠選出健康資料集中潛在的亞健康樣本。

表 5.7 半監督聚類分析各模型之超參數設定

	SDEC
超參數	Max epoch: 50 Batch size: 128 Optimizer: Adam Initial learning rate: 1×10^{-4} Update interval: 5 Stopping threshold: 0.01 Pairwise constraints matrix A: using all sub-healthy data (364 images)

圖 5.5—圖 5.7 可以看到在三個不同模型的分群結果與無監督聚類分析類似，在左側的預測分群中皆可以分成兩個不同的類群，但在右側真實標註分布中卻沒不向預測分群一樣分成兩個類群，也沒有部分健康樣本向亞健康樣本聚集的趨勢，同樣在此聚類分析結果圖中並沒有辦法幫助挑選出潛在的亞健康樣本。故同樣需要透過將此半聚類分析執行十次，統計標註為健康樣本但被模型預測為亞健康的樣本次數，以此方式來嘗試選出潛在的亞健康樣本。

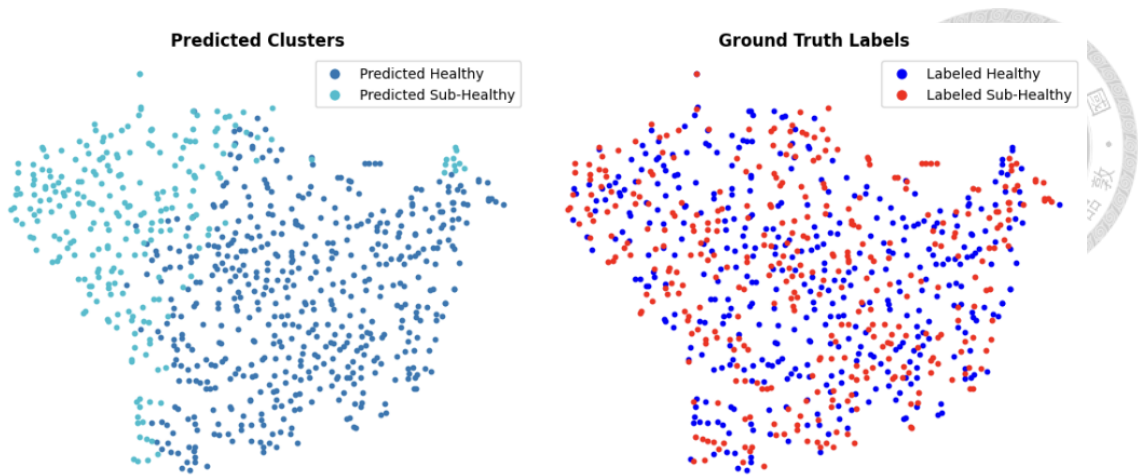


圖 5.5 ImageNet Pre-train Model 半監督聚類分析預測分群與實際分群結果

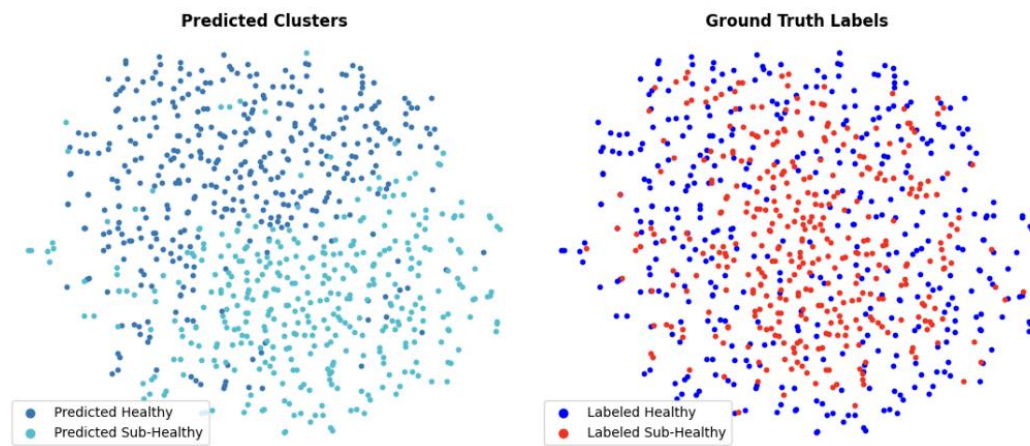


圖 5.6 SimCLR Pre-train Model 半監督聚類分析預測分群與實際分群結果

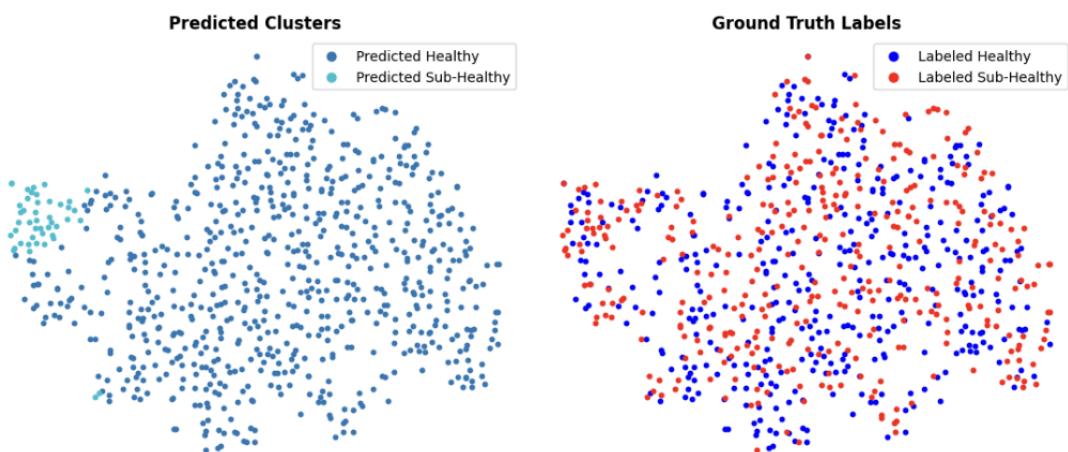


圖 5.7 OCT Pre-train Model 半監督聚類分析預測分群與實際分群結果

在半監督式聚類分析中，ImageNet Pre-train Model 同樣展現出對於預測亞健康樣本的保守傾向。如表 5.7 所示，大多數健康樣本被誤判為亞健康的重複次數並不高，但值得注意的是，有一個樣本在 10 次訓練中被重複預測為亞健康樣本達

9 次。在模型整體預測相對謹慎的情況下，這樣的高度重複性顯得格外顯著，值得懷疑該樣本是否為潛在的亞健康樣本；在 SimCLR Pre-train Model 中，模型同樣在每次分群時從健康資料集中預測出大量的亞健康樣本。然而在表 5.9 中可以看到被重複預測 10 至 8 次的樣本僅有 11 張。這表示儘管模型每次挑選的數量多，但這 11 張樣本在幾乎每一次分群中都被一致地標示為亞健康，顯示其在模型眼中具有明顯的異質性，因此也是值得懷疑該樣本是否為潛在的亞健康樣本；至於 OCT Pre-train Model 在表 5.11 中可看到共有 7 個樣本重複被挑選了十次，8 個樣本重複被挑選了 9 次，此模型對於亞健康樣本的挑選相對保守，卻有 15 個樣本被重複挑選了 9 次甚至 10 次，因此這 15 個樣本同樣值得懷疑次否為潛在的亞健康樣本。

表 5.8 ImageNet Pre-train Model 在 Healthy vs Sub-healthy(Female)資料集半監督聚類分析結果

	1	2	3	4	5	6	7	8	9	10	Avg ±SD
Accuracy(%)	52.5	52.4	49.6	54.8	53.5	48.7	52.5	56.8	49.6	52.7	52.31 ±2.36
Precision(%)	50.2	50	46.8	55.1	52.7	45.8	66.7	54.8	47.2	80	54.93 ±10.14
Recall(%)	33.8	0.3	42.3	28.3	24.2	42.3	0.5	53	47.8	1.1	27.36 ±19.28
F1 score(%)	40.4	0.6	44.4	37.4	33.2	44	1	53.9	47.5	2.2	30.46 ±19.82
Specificity(%)	69.5	99.75	56.25	79	80.25	54.5	99.75	60.25	51.25	99.75	75.025 ±18.61
FPR(%)	30.5	0.25	43.75	21	19.75	45.5	0.25	39.75	48.75	0.25	24.975 ±18.61

表 5.9 ImageNet Pre-train Model 在半監督聚類分析選出潛在亞健康樣本統計結果

Filename	Count
0_00098285_11.png	9
0_00099954_11.png	6
0_00092684_11.png	6
0_00018777_11.png	6
0_00076183_11.png	6
0_00125212_11.png	6

表 5.10 SimCLR Pre-train Model 在 Healthy vs Sub-healthy(Female)資料集半監督聚類分析結果

	1	2	3	4	5	6	7	8	9	10	Avg ±SD
Accuracy(%)	56.4	49.6	47.9	54.2	50.7	48.2	48.4	48	53.1	49	50.55 ±2.85
Precision(%)	53.9	47	45.3	51.8	48.2	45.5	45.5	45.7	50.8	46.4	48.01 ±2.93
Recall(%)	59.3	45.9	45.3	56.3	48.1	44.8	41.5	47.8	52.5	46.2	48.77 ±5.28
F1 score(%)	56.5	46.4	45.3	54	48.1	45.1	43.4	46.7	51.6	46.3	48.34 ±4.05
Specificity(%)	53.75	53	50.25	52.25	53	51.25	54.75	48.25	53.75	51.5	52.175 ±1.83
FPR(%)	46.25	47	49.75	47.75	47	48.75	45.25	51.75	46.25	48.5	47.825 ±1.83

表 5.11 SimCLR Pre-train Model 在半監督聚類分析選出潛在亞健康樣本統計結果

Filename	Count	Filename	Count
0_00084066_11.png	10	0_00092684_11.png	8
0_00085895_11.png	9	0_00081481_11.png	8
0_00078172_11.png	9	0_00027696_11.png	8
0_00061363_11.png	9	0_00087466_11.png	8
0_00122977_11.png	9	0_00110461_11.png	8
0_00050800_11.png	9		

表 5.12 OCT Pre-train Model 在 Healthy vs Sub-healthy(Female)資料集半監督聚類分析結果

	1	2	3	4	5	6	7	8	9	10	Avg ±SD
Accuracy(%)	53.5	52.5	53.3	53.3	53.3	53.3	53.5	53.3	53.1	53.3	53.24 ±0.27

Precision(%)	60.5	53.3	59.5	59	57.8	57.4	60.5	57.8	58.3	57.8	58.19 ±1.95
Recall(%)	7.1	2.2	6	6.3	7.1	7.4	7.1	7.1	5.8	7.1	6.32 ±1.47
F1 score(%)	12.7	4.2	10.9	11.4	12.6	13.1	12.7	12.6	10.6	12.6	11.34 ±2.51
Specificity(%)	95.75	98.25	96.25	96	95.25	95	95.75	95.25	96.25	95.25	95.9 ±0.89
FPR(%)	4.25	1.75	3.75	4	4.75	5	4.25	4.75	3.75	4.75	4.1 ±0.89

表 5.13 OCT Pre-train Model 在半監督聚類分析選出潛在亞健康樣本統計結果

Filename	Count	Filename	Count
0_00086033_11.png	10	0_00100240_11.png	9
0_00085952_11.png	10	0_00071975_11.png	9
0_00040840_11.png	10	0_00026190_11.png	9
0_00037621_11.png	10	0_00049489_11.png	9
0_00087466_11.png	10	0_00099954_11.png	9
0_00098285_11.png	10	0_00096033_11.png	9
0_00087183_11.png	10	0_00035869_11.png	9
0_00019636_11.png	9		

5.4 聚類結果之分析與討論

J. Donald M. Gass 在 1987 年把黃斑前膜畫分成三個等級: [37]

1. 第零級(Grade 0):在此級的特徵為透明的薄膜，還未使視網膜變形或遮蔽下方的血管。可能有不規則的反光，但邊緣還不太明顯。
2. 第一級(Grade 1):在此級已經使視網膜產生不規則的皺褶，但下方的血管還可見，可以觀察到黃斑部周圍的血管被拉扯向中心收縮。
3. 第二級(Grade 2):在此級下方的血管已經被遮蔽，並有明顯的視網膜變形及弓狀血管聚集，常導致中央凹異位(Foveal Ectopia)、異位排列(Heterotopia)或淺層桌面型視網膜剝離(Tabletop Retinal Detachment)。

目前臨床上對「黃斑前膜之亞健康狀態」尚無明確的診斷依據，僅能仰賴雙眼影像進行對照。根據臨床經驗，亞健康影像的特徵應介於健康與第零級黃斑前膜之間，但由於缺乏明顯病灶，無法直接進行標註與分類。因此在本實驗中是希望能

從健康資料集中進一步識別出可能具有亞健康特徵的潛在樣本。

如上所述，本實驗主要目的並不是像一般的分類任務追求高分類準確率，而是要從單一類別(健康樣本)中挑出可能標註錯誤的樣本(亞健康樣本)，並且在本實驗中對於亞健康影像之標註較有信心，因此對於優秀模型的期待也會與一般分類任務不同，首先期待模型分群時在預測的亞健康樣本中有盡可能多來自亞健康資料集，也就是希望精確率可以越高越好，再來希望模型在預測亞健康樣本時可以較為保守，而不是每次的分群都預測大量的影像為亞健康樣本，也就是特異度越高越好而假陽性率越低越好，所以在本實驗中所重視的數據會是精確率、特異度、假陽性率這三項。OCT Pre-train Model 不只在預測的影像上有高度的一致性，並且服合本實驗對於模型的期待，因此這 15 張影像是認為最有可能的潛在亞健康樣本。除此之外還對這些可能的潛在亞健康樣本進行 Grad-CAM 分析，分別使用先前在健康與亞健康之分類任務中訓練的三種模型來觀察對於這些影像模型的觀察重點會在哪些區域。

在將這些影像與謝易庭醫師討論過後，謝醫師表示亞健康影像與正常健康影像在中央凹及其附近的區域之厚薄度會有些許的不同，除此之外中央凹的弧度相對於正常健康影像也較為平緩，如圖 5.8 所示紅色圈起來的部分為亞健康影像與健康影像會有些許差異之地方。在圖 5.9 中可以看到 ImageNet Pre-train Model 在經過健康與亞健康分類訓練後對於潛在亞健康樣本觀察重點落在中央凹之兩側，與謝醫師認為重要之區域吻合；在圖 5.10 中可以看到 SimCLR Pre-train Model 在經過健康與亞健康分類訓練後對於潛在亞健康樣本觀察重點落在中央凹及左側之下方，雖然觀察區域偏下但謝醫師認為仍有觀察到重點區域；在圖 5.11 中可以看到 OCT Pre-train Model 在經過健康與亞健康分類訓練後對於潛在亞健康樣本觀察重點落在中央凹左側及視網膜色素上皮(Retinal Pigment Epithelium; RPE)下方，謝醫師認為中央凹左側的部分與觀察重點吻合，但目前研究上對於視網膜色素上皮下方的區域與亞健康影像之關聯性還未發現。總結來說，謝醫師認為三種不同模型經過建網與亞健康分類訓練後，雖然 OCT Pre-train Model 有一部分重點觀察在視網膜色素上皮下方的區域，但整體上模型觀察之重點區域皆合理。

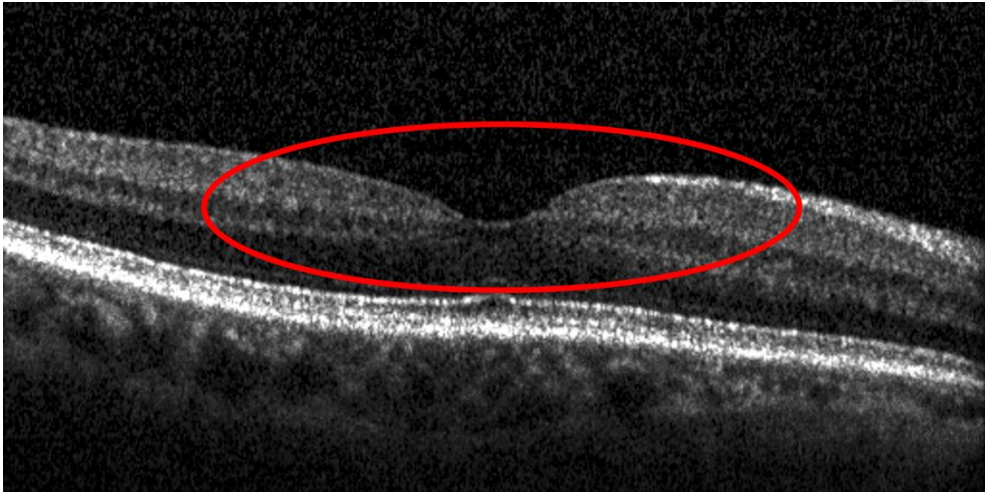


圖 5.8 潛在亞健康樣本之影像

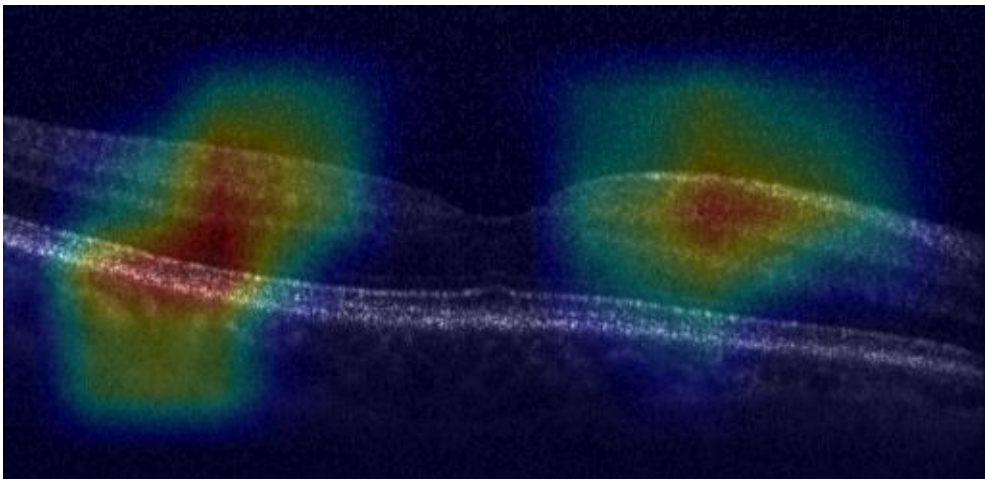


圖 5.9 ImageNet Pre-train Model 經健康與亞健康分類訓練後對潛在亞健康樣本之 Grad-CAM 影像(0_00092684_11.png)

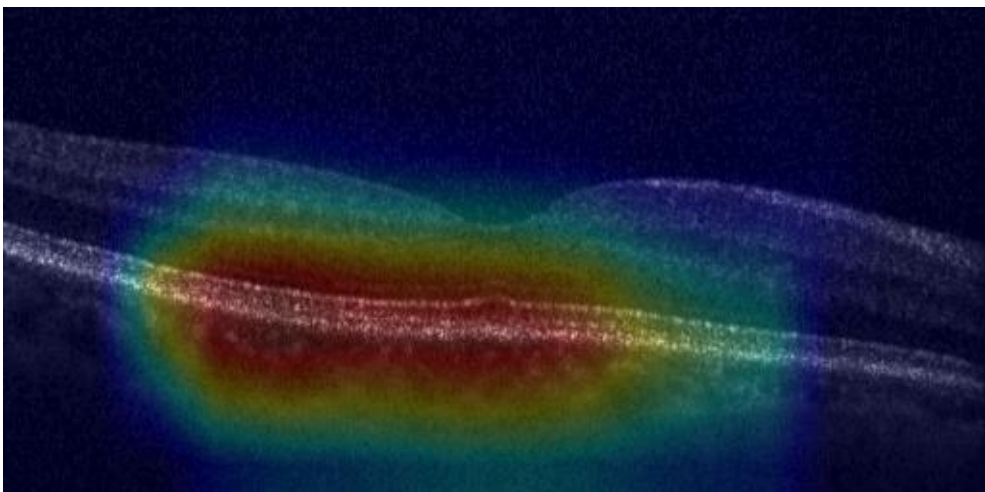


圖 5.10 SimCLR Pre-train Model 經健康與亞健康分類訓練後對潛在亞健康樣本

之 Grad-CAM 影像(0_00092684_11.png)

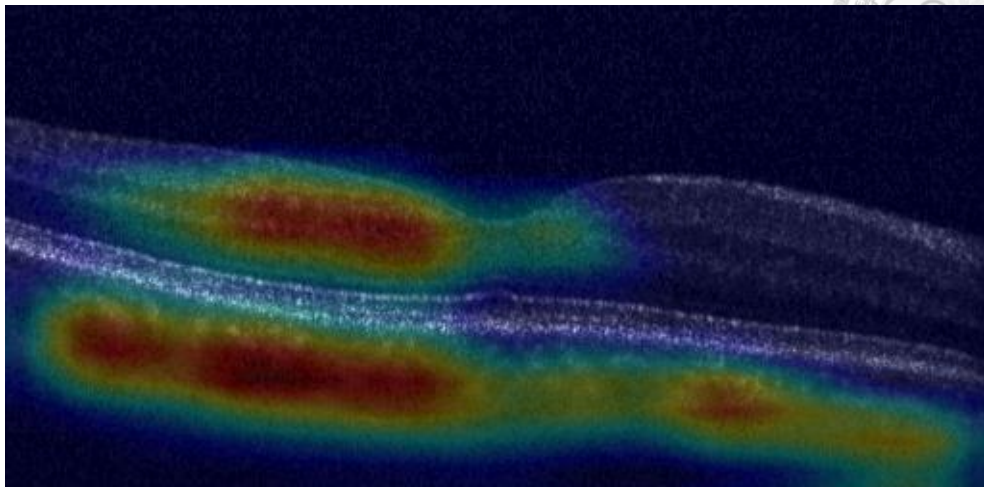


圖 5.11 OCT Pre-train Model 經健康與亞健康分類訓練後對潛在亞健康樣本之 Grad-CAM 影像(0_00092684_11.png)

觀察 K-means 演算法與 SDEC 演算法對潛在亞健康樣本統計結果可以看到，K-means 之統計結果在 ImageNet Pre-train Model 中影像被重複挑選的次數過低；在 SimCLR Pre-train Model 與 OCT Pre-train Model 雖然影像被重複挑選的次數較高，但因為在每次的分群中都選出大量的影像預測為亞健康，因此最終選出之潛在亞健康影像也相當大量。SDEC 之統計結果在 ImageNet Pre-train Model 中與 K-means 演算法之統計結果相似，但有一張影像被模型重複挑選了九次；在 SimCLR Pre-train Model 中雖然在每次的分群中都會選出大量的影像預測為亞健康，但觀察被重複預測 10 至 8 次的樣本僅有 11 張；在 OCT Pre-train Model 中則是在十次的分群中皆選擇了相同的 18 張影像預測為亞健康樣本。

監督式聚類與半監督式聚類的主要差異，在於是否引入配對限制指示矩陣 (Pairwise Constraint Matrix)。此矩陣用以表達樣本之間的相對關係，其中的數值包含 1、0 與 -1，分別代表兩樣本標註相同、無標註資訊或標註不同。當兩個樣本皆被標註且屬於同一類別時，矩陣對應位置為 1，表示模型應鼓勵該兩樣本在特徵空間中靠近；若兩者標註不同則為 -1，模型應施加相反的約束，避免其被分入同一群。而當至少一方未被標註時，對應位置則為 0，表示該樣本對在計算目標函數時配對損失函數會被忽略。在本研究中，由於僅提供亞健康樣本之標註資訊，未對健康樣本進行標註，因此建構出的配對限制指示矩陣中僅包含 1 與 0 兩種數值，如圖 5.12 所示。雖然亞健康影像在整體資料集中佔比約 47.6%，然而 SDEC 所關注的是兩兩樣本之間的相對關係，故可構成有效 must-link 約束的樣本對數量遠低於樣本標註比例。經計算，整體 764x764 的矩陣中僅有 11.38% 的元素為 1，代表具有明確相似性資訊的樣本對比例相對稀少。

儘管配對資訊在數量上顯得有限，實驗結果顯示這樣的約束仍對模型訓練具有

實質幫助。推測原因在於這些明確的同群關係即使數量不多，依然能有效引導模型學習潛在的語意結構，從而提升聚類之穩定性。若未來能進一步增加亞健康樣本數，預期能提高配對矩陣中具資訊樣本對的比例，進而進一步改善模型聚類品質。

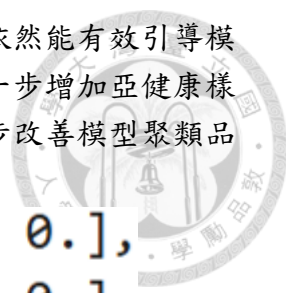

$$\begin{bmatrix} [0., 0., 0., \dots, 0., 0., 0.], \\ [0., 0., 0., \dots, 0., 0., 0.], \\ [0., 0., 0., \dots, 0., 0., 0.], \\ \dots, \\ [0., 0., 0., \dots, 0., 1., 1.], \\ [0., 0., 0., \dots, 1., 0., 1.], \\ [0., 0., 0., \dots, 1., 1., 0.] \end{bmatrix}$$

圖 5.12 SDEC 訓練中之配對限制指示矩陣

Chapter 6 結果與未來展望



6.1 結論

在本實驗中使用不同深度學習模型及機器學習方式探討對於亞健康影像分辨之可行性，主要可以分為兩個部分。

在第一部分中首先嘗試使用自監督學習中的 SimCLR 演算法用約 20000 張之視網膜 OCT 影像訓練 ResNet-18 模型並與使用 ImageNet 預訓練之 ResNet-18 在相對簡單的健康與不健康分類任務、黃斑前膜與其他疾病分類任務、健康與黃斑前膜分類任務中進行比較，從結果中可以看到兩種模型在各項任務中皆可達到 80% 以上之準確率，皆可以分辨該項任務中的兩類影像，但 ImageNet Pre-train Model 之準確率普遍高於 SimCLR Pre-train Model。認為原因是因為用來做 SimCLR 預訓練之資料量不足，在其他研究中指出因為自監督訓練之標註是演算法自行產生，因此模型可能不會按照我們的想法學會如何看一張視網膜 OCT 影像，而是偏好學習一些「捷徑」像亮度、對比度、紋理等來學習，因為缺乏人為引導模型學習因此需要更為龐大的資料集讓模型能夠真正學會看視網膜 OCT 影像。這也與本實驗之結果吻合，在 Grad-CAM 影像中可以觀察到，SimCLR Pre-train Model 在判讀影像時，常會將注意力集中於亮度較高的區域，例如神經纖維層 (Nerve Fiber Layer)，即使這些區域並非該類影像的關鍵特徵，模型仍可能重點觀察此區域。而在健康與黃斑前膜分類任務中，因為分別兩類影像之重點為內界膜 (Inner Limiting Membrane) 上的亮帶部分，恰巧此部分本就是 SimCLR Pre-train Model 會視為捷徑並重點觀察的部分，因此此任務也是唯一一項 SimCLR Pre-train Model 表現優於 ImageNet Pre-train Model 的任務。接下來將這兩種模型與先前研究中使用之 OCT Pre-train Model 在健康與亞健康分類任務中進行比較，在此任務中三個模型雖有些許差異但準確率皆落在 66% 左右，在此項任務中的差異與前三項任務相比明顯較小，因此懷疑資料集中是否有標註上的錯誤。

在亞健康眼睛需要對側眼已確診為黃斑前膜才能進行標註，而我們懷疑在健康影像中有部分影像已經是亞健康之狀態，但對側眼還未發病因此無法標註為亞健康，在第二部分的實驗中試圖找出健康資料集中潛在的亞健康影像。首先使用無監督聚類分析中的 K-means 演算法，但三種模型在視覺化結果中實際標註結果皆沒規則的散布，不像預測結果一樣分成兩群。在統計重複挑選潛在亞健康影像之結果中，ImageNet Pre-train Model 重複挑選之潛在亞健康樣本次數偏低；SimCLR Pre-train Model 與 OCT Pre-train Model 雖然重複挑選之潛在亞健康樣本次數可以達到 10 次至 8 次，但選出的影像數量也都超過 40 張，認為超過 10% 健康資料集影像標註有問題，因此認為以上這三種模型使用 k-means 演算法無法

根據視覺化結果從健康資料集中挑選出潛在的亞健康樣本，而在潛在亞健康樣本之統計結果中代表性與一致性皆不足，同樣難以作為合理懷疑樣本標註錯誤的依據。

在半監督聚類分析中使用 SDEC 演算法，此種演算法之優點為可以只告訴模型其中一個類別的標註，在本實驗中僅有亞健康樣本的標註有明確的定義，因此透過只告訴模型亞健康樣本之標註來引導模型學習亞健康樣本之特徵，並藉此從健康資料集中挑選出潛在的亞健康樣本。在三種模型的視覺化結果中同樣顯示出預測結果分為兩群但實際標註結果為沒規則之散布。但在統計重複挑選潛在亞健康影像之結果中，ImageNet Pre-train Model 在每次預測中挑選之潛在亞健康樣本數量不多，但卻有一張影像在十次中皆被重複預測為亞健康樣本；SimCLR Pre-train Model 在每次預測中雖然挑選了大量的潛在亞健康樣本，但其中僅有 11 張影像被重複挑選了 10 至 8 次；OCT Pre-train Model 儘管每次預測中挑選潛在亞健康樣本較為保守，仍然有 15 張影像被重複挑選了 10 至 9 次。在將這些影像與謝易庭醫師討論後，謝醫師認為這些影像之中央凹弧度確實相對較平緩，與先前研究中對於亞健康影像之觀察符合，因此透過 SDEC 演算法幫助從健康資料集中選出部分值得懷疑是否真的標註錯誤的潛在亞健康樣本。

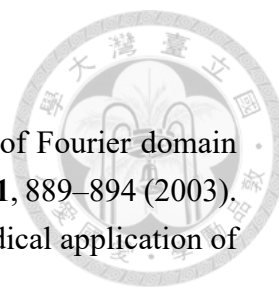
在本研究中發現有經過視網膜 OCT 影像微調之 OCT Pre-train Model 不論是在監督式分類任務或半監督聚類分析中表現皆最優秀，在健康與亞健康分類任務中雖差異不大但仍然有著較高的準確率與較小的標準差，在半監督聚類分析中有著最高的精確率(Precision)，代表在預測為亞健康之影像中大部分影像確實為亞健康樣本，且在預測亞健康樣本保守的情況下依然穩定的重複選取 15 張影像。雖然 ImageNet Pre-train Model 與 SimCLR Pre-train Model 均僅經歷預訓練階段，未進行與下游任務相關的微調，但在多數監督式分類任務中，ImageNet Pre-train Model 的表現明顯優於 SimCLR Pre-train Model。儘管在預訓練階段 SimCLR Pre-train Model 使用之資料與後續任務同為視網膜 OCT 影像而 ImageNet Pre-train Model 則為一般自然影像，兩者在資料型態上有所差異，但因為數量上的差異使得僅使用約 20000 張影像之 SimCLR Pre-train Model 無法真正學會如何「理解」影像內容，反而傾向倚賴影像中的低階特徵或捷徑來完成任務，ImageNet Pre-train Model 則透過約 1280 萬張影像使其在特徵學習過程中更能掌握真實的語意結構與視覺判讀方式，進而在後續的分類任務中展現出更穩定且優異的表現。

6.2 未來展望

在本實驗中嘗試使用自監督學習之 SimCLR 演算法預訓練模型，雖然在最終結果中不如 ImageNet Pre-train Model，然而這並不代表 SimCLR 本身無效，希望能夠透過增加資料集規模改善模型對於捷徑之依賴，進而在下游任務中減少人工標註的時間與精力卻能達到相同甚至是更好的結果。除此之外因為 ImageNet 資料

集中均為一般自然影像與醫學影像仍有差距，而使用 SimCLR 演算法則可以根據下游任務使用相同類型之影像進行預訓練，使模型能夠學習更目標任務的影像分布，進而提升其特徵學習的相關性與應用性。

在本實驗中初步選出一些值得懷疑的潛在亞健康樣本，若能持續追蹤這些病患後續是否發展為黃斑前膜，將有助於驗證本研究方法在早期風險篩檢中的實用性。透過長期觀察這些樣本的病程變化，亦有機會進一步發掘潛在亞健康影像中尚未被識別的重要結構特徵，從而豐富對亞健康狀態的理解。期望未來此類方法能在臨床上發揮預防性診斷的功能，協助及早辨識高風險個案，進而能夠提前預防。



Reference

1. R. Leitgeb, C. Hitzenberger, and A. F. Fercher, "Performance of Fourier domain vs. time domain optical coherence tomography," *Opt. Express* **11**, 889–894 (2003).
2. A. Chiang and C. Y. Chang, "Technology overview and biomedical application of optical coherence," *MaterialsNet* **370**, 108–116 (2017).
3. B. E. Bouma, S.-H. Yun, B. J. Vakoc *et al.*, "Fourier-domain optical coherence tomography: recent advances toward clinical utility," *Curr. Opin. Biotechnol.* **20**, 111–118 (2009).
4. K. H. Nguyen, B. C. Patel, and P. Tadi, "Anatomy, head and neck: Eye retina," in *StatPearls* [Internet], StatPearls Publishing (2023).
5. U. of Tennessee Health Science Center, "Layers of the Retina," *Hamilton Eye Institute* (2022).
6. Discovery Eye Foundation, "Layers of the Retina," <https://discoveryeye.org/layers-of-the-retina/>. [Accessed: May 26, 2025].
7. V. Mansouri, "X-linked retinitis pigmentosa gene therapy: preclinical aspects," *Ophthalmol. Ther.* **12**, 7–34 (2023).
8. K. A. Hussey, S. E. Hadyniak, and R. J. Johnston Jr, "Patterning and development of photoreceptors in the human retina," *Front. Cell Dev. Biol.* **10**, 878350 (2022).
9. R. S. Molday and O. L. Moritz, "Photoreceptors at a glance," *J. Cell Sci.* **128**, 4039–4045 (2015).
10. B. A. Labib, "Macular Focus," <https://www.reviewofoptometry.com/article/macular-focus>. [Accessed: May 27, 2025].
11. J. M. Provis, P. L. Penfold, E. E. Cornish *et al.*, "Anatomy and development of the macula: specialisation and the vulnerability to macular degeneration," *Clin. Exp. Optom.* **88**, 269–281 (2005).
12. I. Rehman, N. Mahabadi, M. Motlagh *et al.*, "Anatomy, head and neck, eye fovea," in *StatPearls* [Internet], StatPearls Publishing (2023).
13. J. S. Schiffman, N. B. Patel, R. A. Cruz *et al.*, "Optical coherence tomography for the radiologist," *Neuroimaging Clin. N. Am.* **25**, 367–382 (2015).
14. A. T. Fung, J. Galvin, and T. Tran, "Epiretinal membrane: a review," *Clin. Exp. Ophthalmol.* **49**, 289–308 (2021).
15. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE* **86**, 2278–2324 (2002).
16. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Adv. Neural Inf. Process. Syst.* **25**, (2012).
17. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 770–778 (2016).

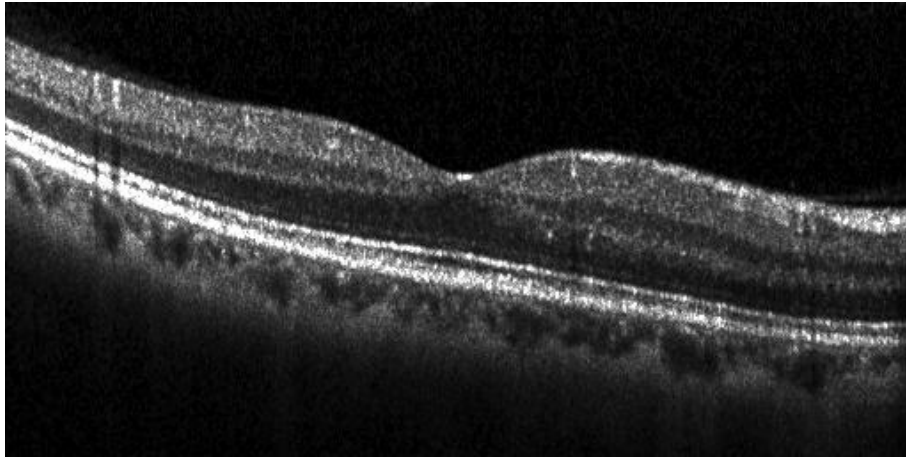
18. C. A. Taiwan, "深度學習：CNN 原理," <https://cinnamonaitaiwan.medium.com/%E6%B7%B1%E5%BA%A6%E5%AD%B8%E7%BF%92-cnn%E5%8E%9F%E7%90%86-keras%E5%AF%A6%E7%8F%BE-432fd9ea4935> (accessed 15 Jun. 2025).
19. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* 521, 436–444 (2015).
20. V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proc. 27th Int. Conf. Mach. Learn. (ICML-10)*, 807–814 (2010).
21. Gopalakrishna Adusumilli, "Activation Function- ReLU," <https://gkadusumilli.github.io/relu/> (accessed 18 Jun. 2025).
22. R. Singh, "Decoding CNNs: A Beginner's Guide to Convolutional Neural Networks and their Applications," <https://ravjot03.medium.com/decoding-cnns-a-beginners-guide-to-convolutional-neural-networks-and-their-applications-1a8806cbf536> (accessed 15 Jun. 2025).
23. C. M. Lee, "直觀理解 ResNet — 簡介、觀念及實作(Python Keras)," <https://medium.com/@rossleecooloh/%E7%9B%B4%E8%A7%80%E7%90%86%E8%A7%A3resnet-%E7%B0%A1%E4%BB%8B-%E8%A7%80%E5%BF%B5%E5%8F%8A%E5%AF%A6%E4%BD%9C-python-keras-8d1e2e057de2> (accessed 15 Jun. 2025).
24. C. A. Taiwan, "損失函數的設計 (Loss Function)," <https://cinnamonaitaiwan.medium.com/cnn%E6%A8%A1%E5%9E%8B-%E6%90%8D%E5%A4%B1%E5%87%BD%E6%95%B8-loss-function-647e13956c50> (accessed 15 Jun. 2025).
25. T. Huang, "機器/深度學習：基礎介紹-損失函數(loss function)," <https://chih-sheng-huang821.medium.com/%E6%A9%9F%E5%99%A8-%E6%B7%B1%E5%BA%A6%E5%AD%B8%E7%BF%92-%E5%9F%BA%E7%A4%8E%E4%B8%B8%E7%B4%B9-%E6%90%8D%E5%A4%B1%E5%87%BD%E6%95%B8-loss-function-2dcac5ebb6cb> (accessed 15 Jun. 2025).
26. D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature* 323, 533–536 (1986).
27. X. Liu, et al., "Self-supervised learning: Generative or contrastive," *IEEE Trans. Knowl. Data Eng.* 35, 857–876 (2021).
28. T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in *Int. Conf. Mach. Learn.*, 1597–1607 (2020).
29. J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Math. Stat. Probab.*, 5, 281–298 (1967).
30. C. C. Aggarwal, A. Hinneburg, and D. A. Keim, "On the surprising behavior of distance metrics in high dimensional space," in *Int. Conf. Database Theory*, 420–

- 434 (2001).
31. J. Xie, R. Girshick, and A. Farhadi, "Unsupervised deep embedding for clustering analysis," in *Int. Conf. Mach. Learn.*, 478–487 (2016)
 32. dengliu2000, " Day 20 - 卷積神經網路 (CNN) 的介紹," <https://ithelp.ithome.com.tw/articles/10304119> (accessed 18 Jun. 2025).
 33. 李劭允, "利用卷積神經網路從黃斑部光學同調斷層掃描影像分辨健康及好發黃斑前膜的亞健康眼睛," 碩士, 光電工程研究所, 國立臺灣大學, 2023 年.
 34. R. Geirhos, J.-H. Jacobsen, C. Michaelis, R. Zemel, W. Brendel, M. Bethge, and F. A. Wichmann, "Shortcut learning in deep neural networks," *Nat. Mach. Intell.* 2, 665–673 (2020).
 35. S. Azizi, K. Mustafa, F. Xing, L. Angus, H. L. Jin, T. Epstein, K. Roth, and N. Houlsby, "Big self-supervised models advance medical image classification," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 3478–3488 (2021).
 36. Y.-C. Chen, "如何辨別機器學習模型的好壞？秒懂 Confusion Matrix," <https://ycc.idv.tw/confusion-matrix.html> (accessed 10 Jul. 2025).
 37. P. N. Jain, "Epiretinal Membrane," <https://www.eophtha.com/posts/epiretinal-membrane> (accessed 10 Jul. 2025).

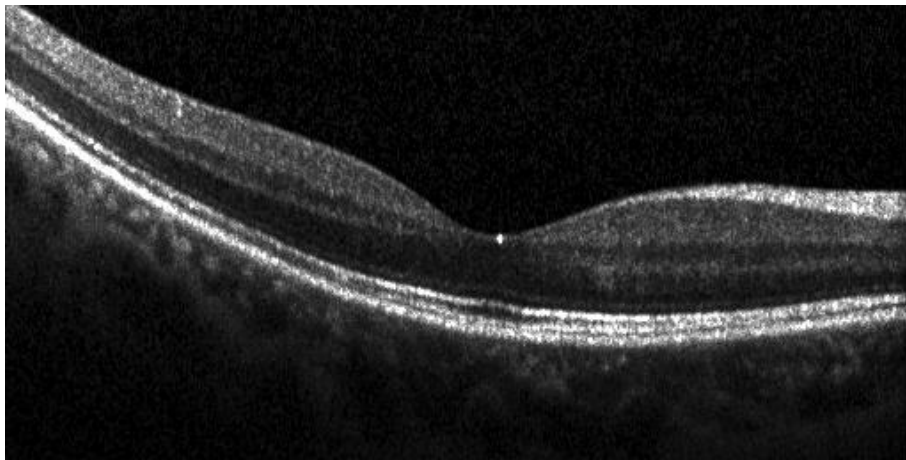
附錄 1 三種不同預訓練模型在 K-means 分群中所選
之潛在亞健康樣本



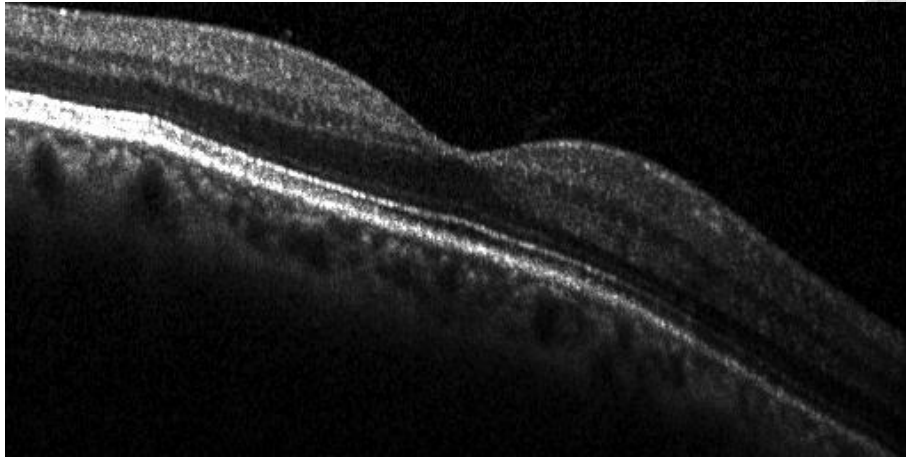
ImageNet Pre-train Model



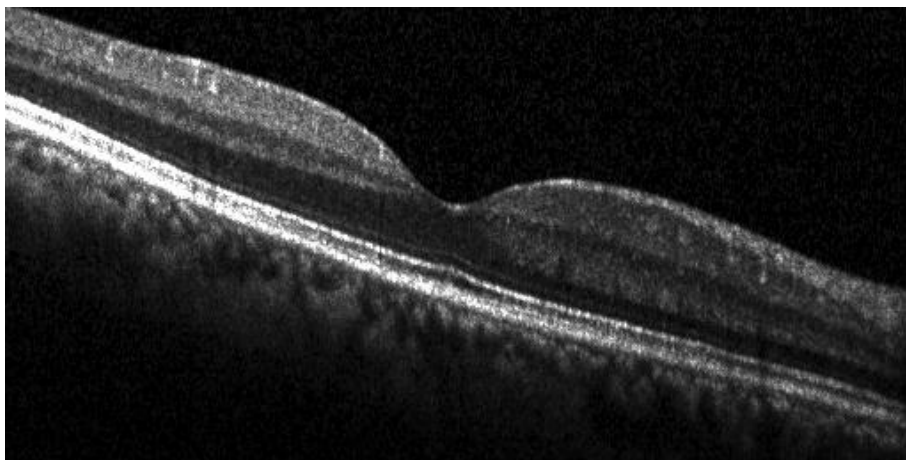
0_00021034_11.png



0_00085015_11.png

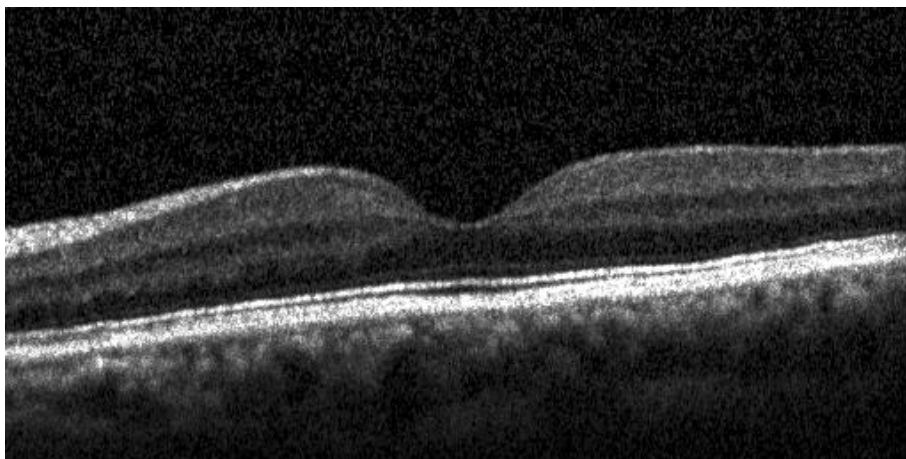


0_00086033_11.png

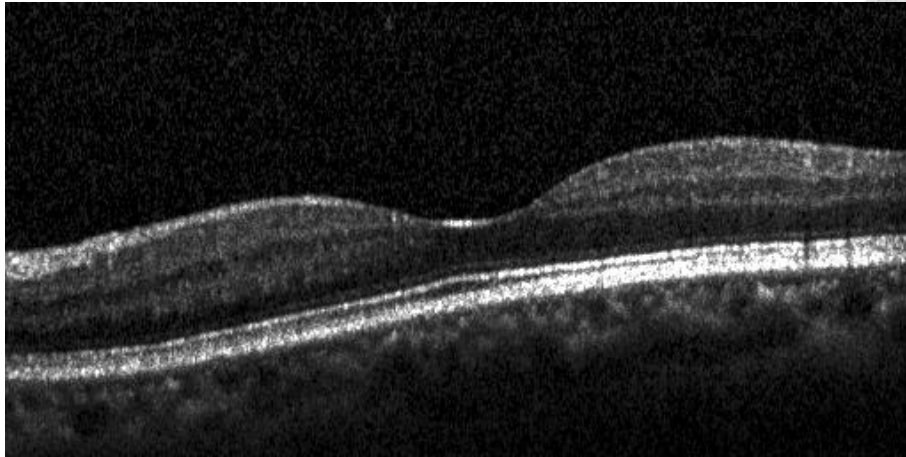


0_00087232_11.png

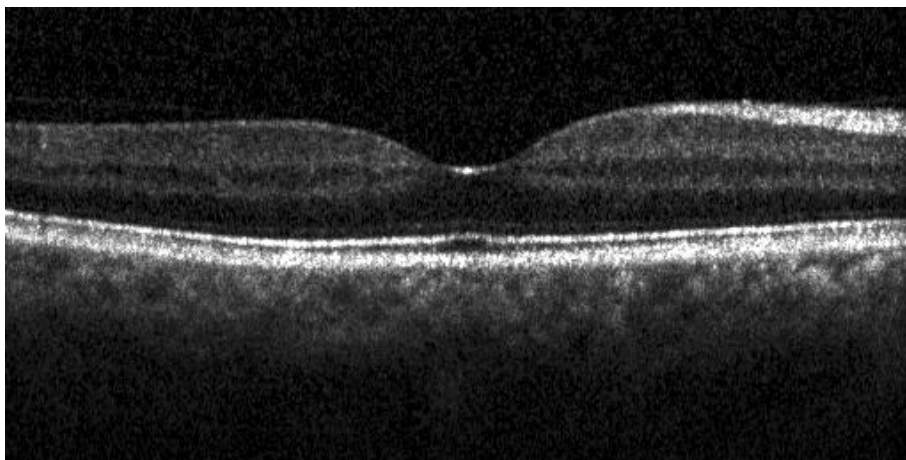
SimCLR Pre-train Model



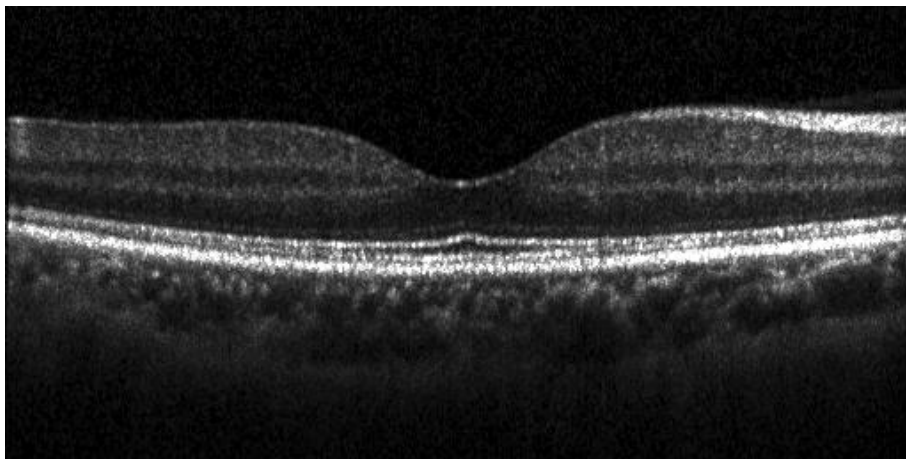
0_00061144_11.png



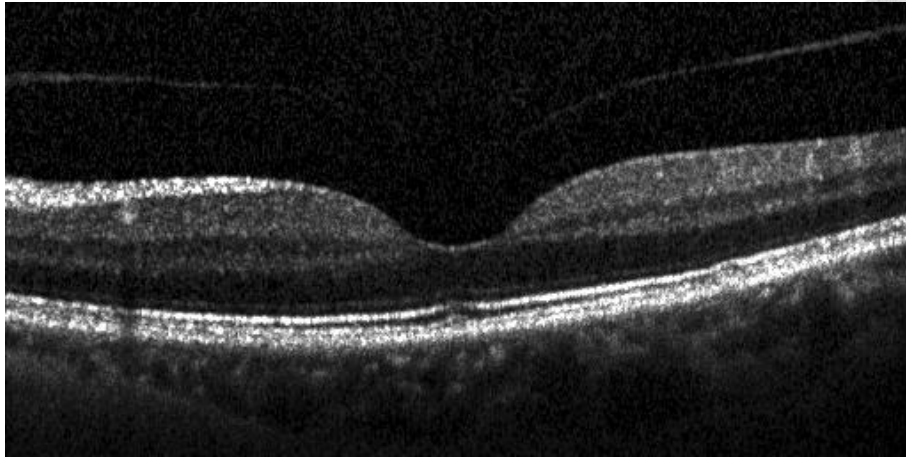
0_00103124_11.png



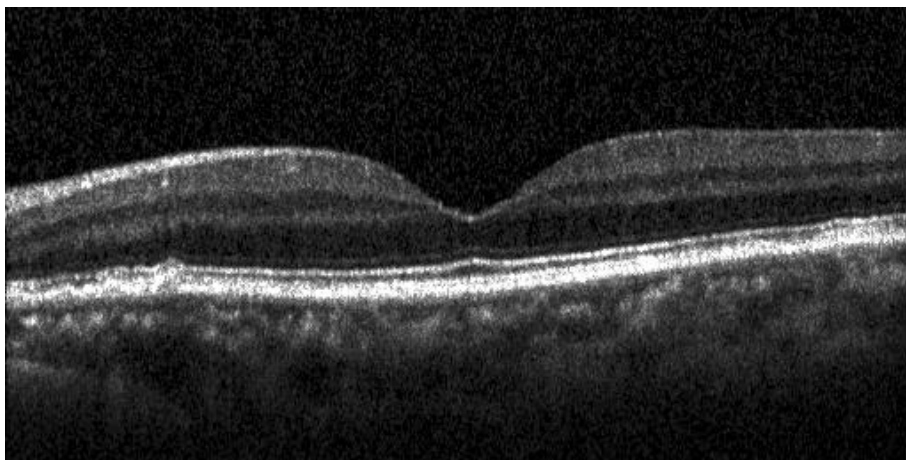
0_00104960_11.png



0_00108438_11.png

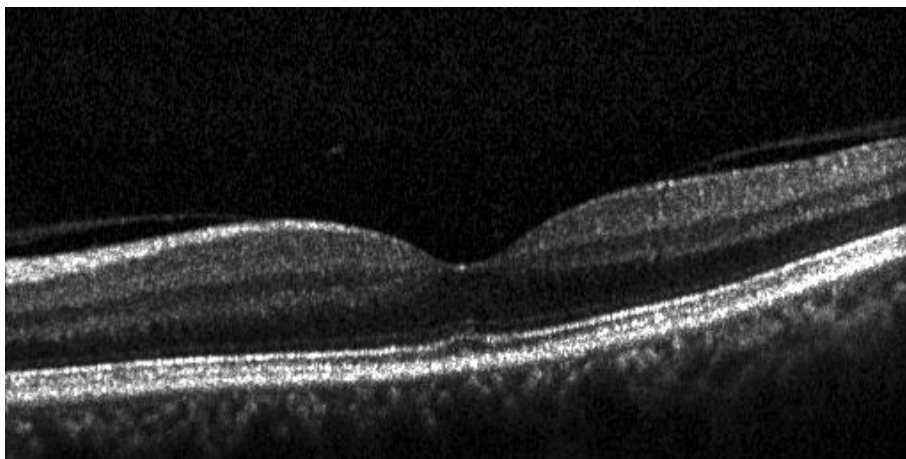


0_00114324_11.png

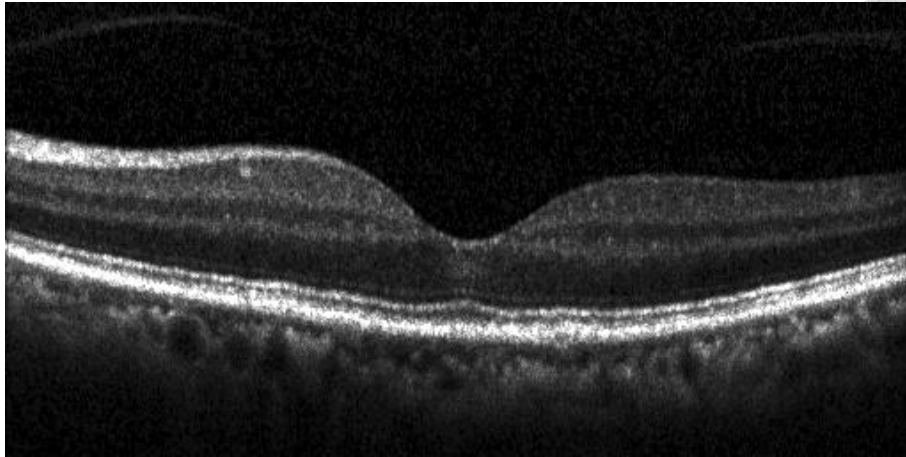


0_00117181_11.png

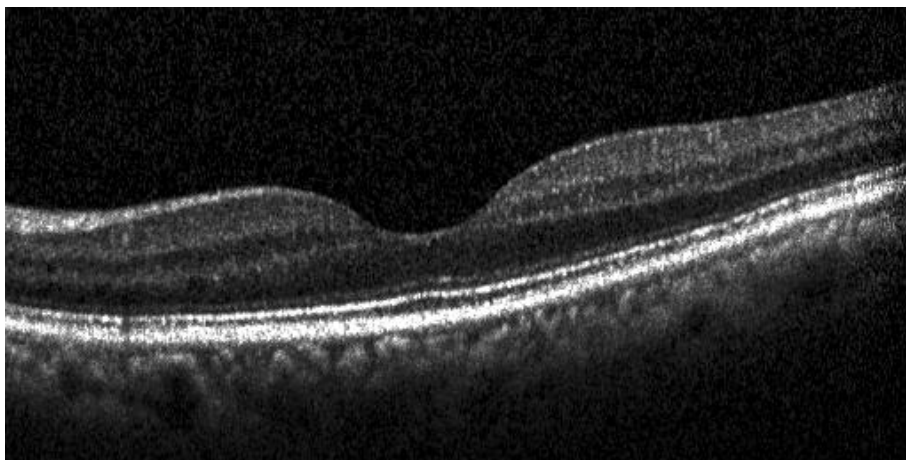
OCT Pre-train Model



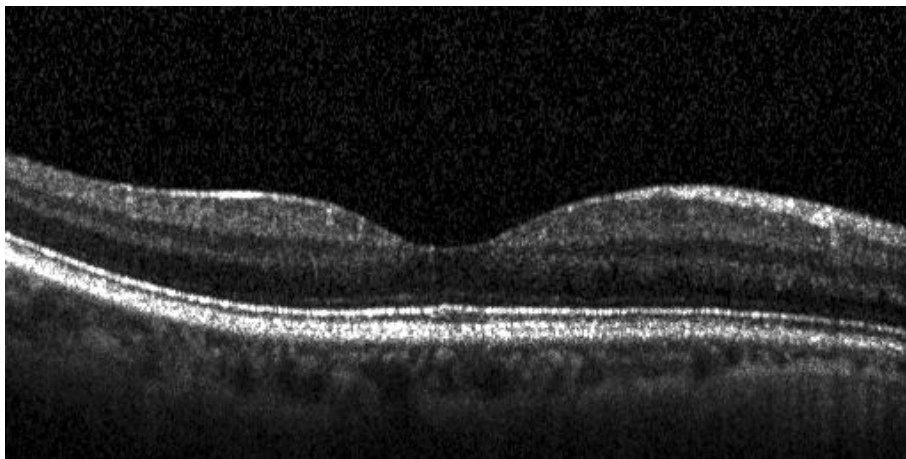
0_00068590_11.png



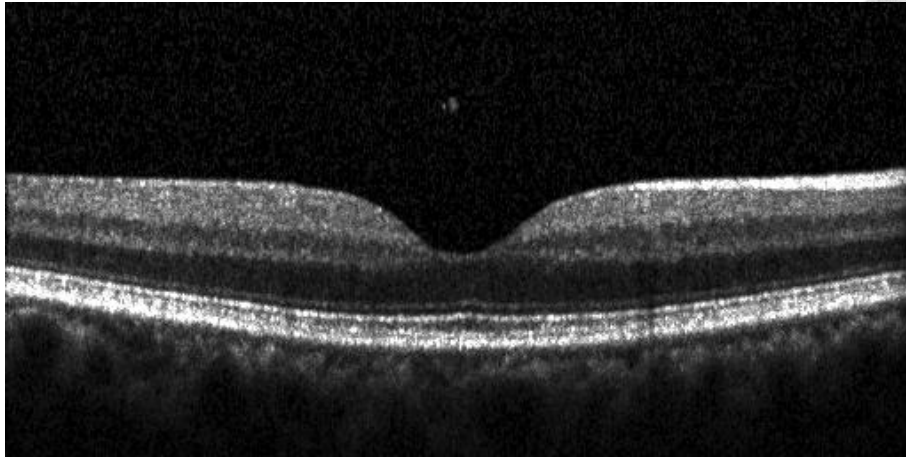
0_00073292_11.png



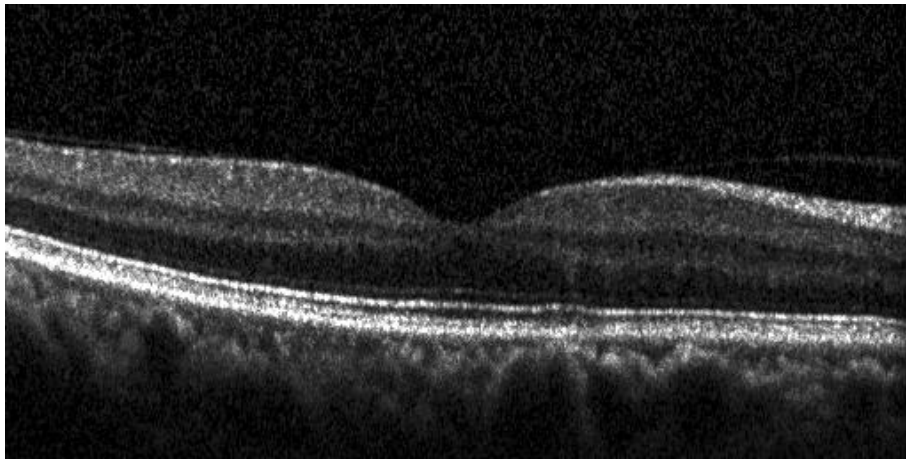
0_00083362_11.png



0_00087204_11.png



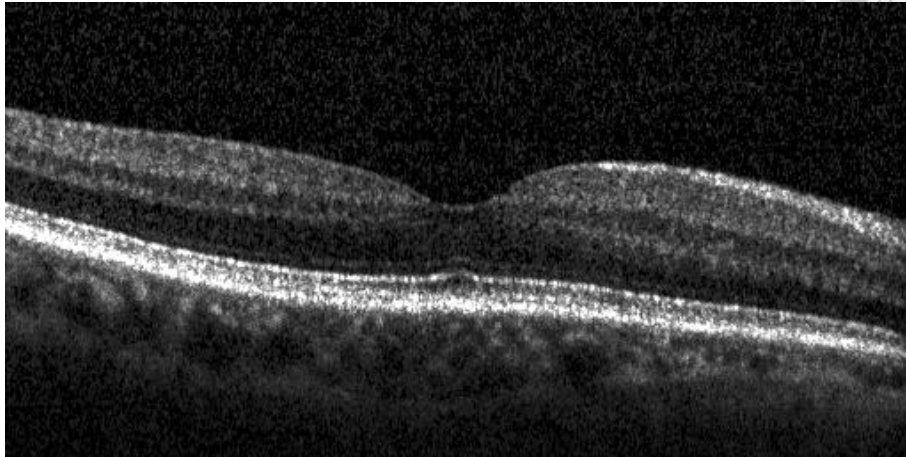
0_00091303_11.png



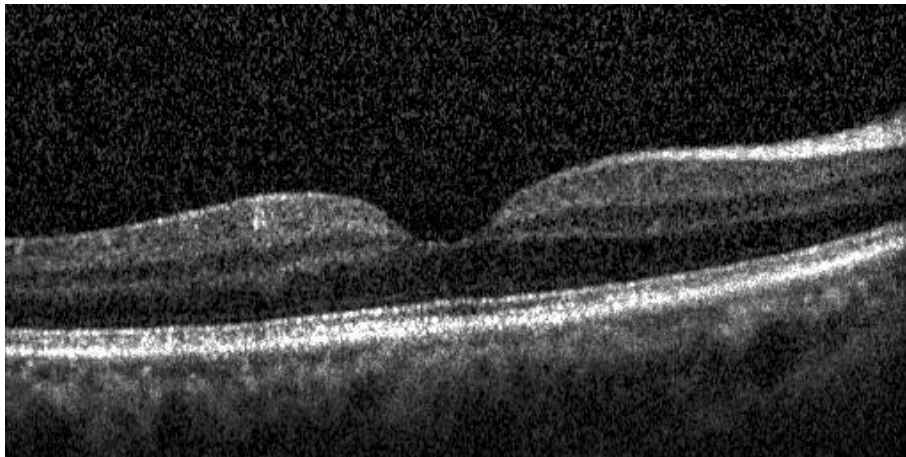
0_00098634_11.png

**附錄 2 三種不同預訓練模型在 SDEC 分群中所選之
潛在亞健康樣本**

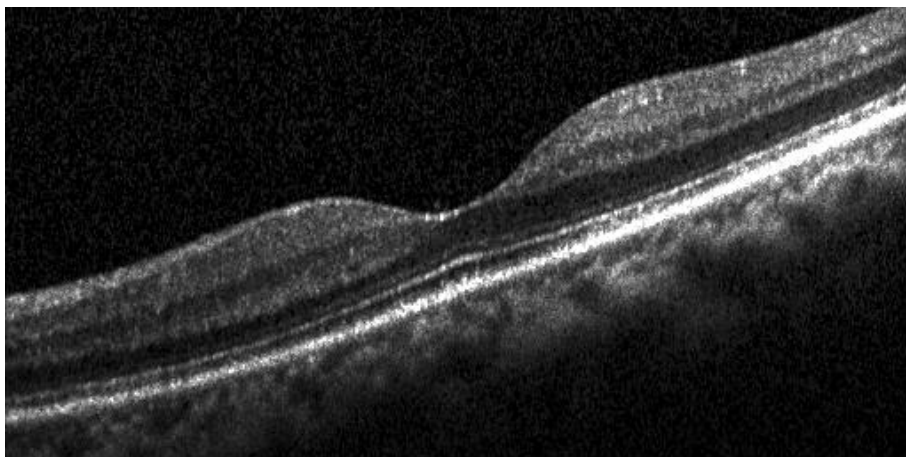
ImageNet Pre-train Model



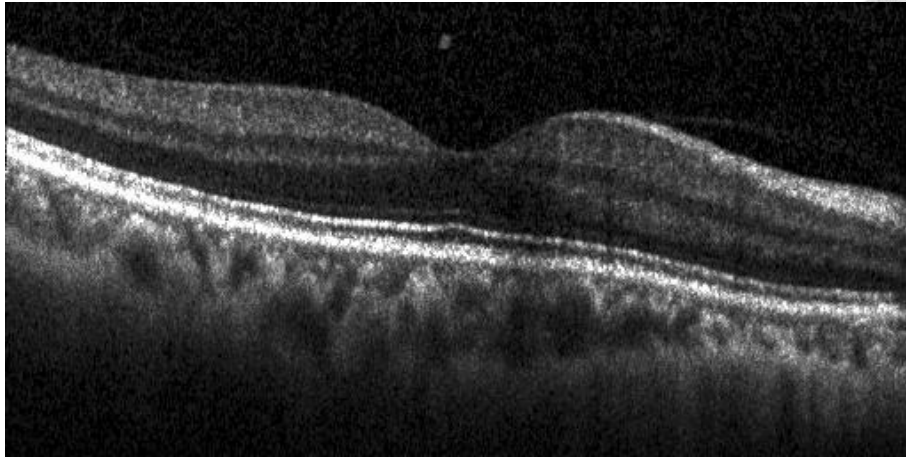
0_00092684_11.png



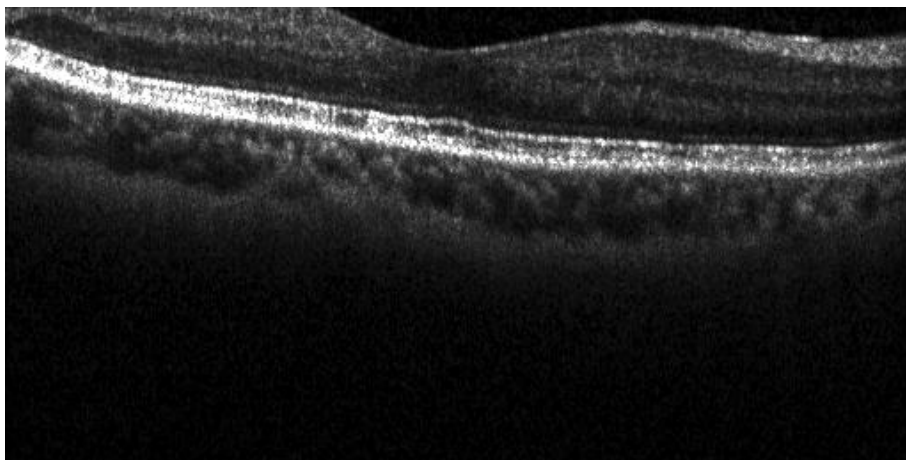
0_00125212_11.png



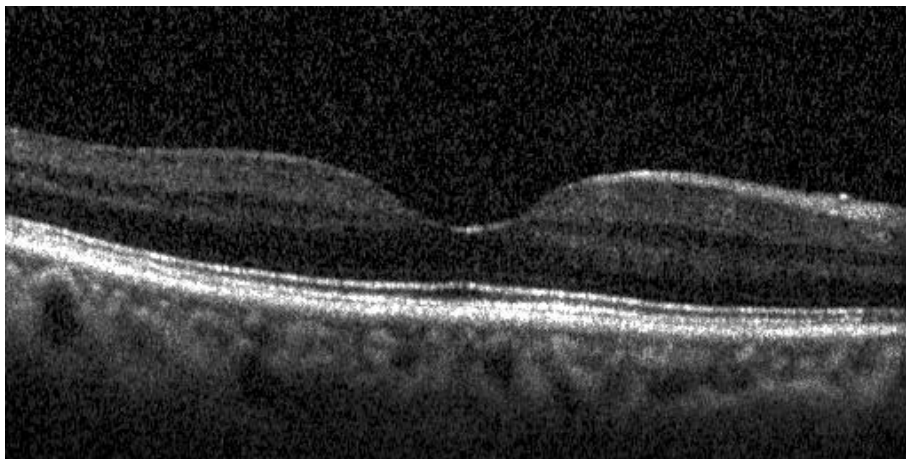
0_00099954_11.png



0_00018777_11.png

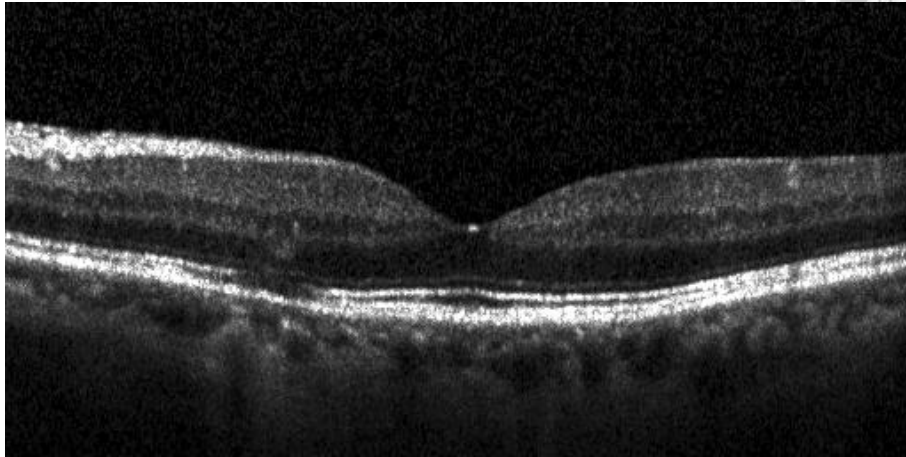


0_00098285_11.png

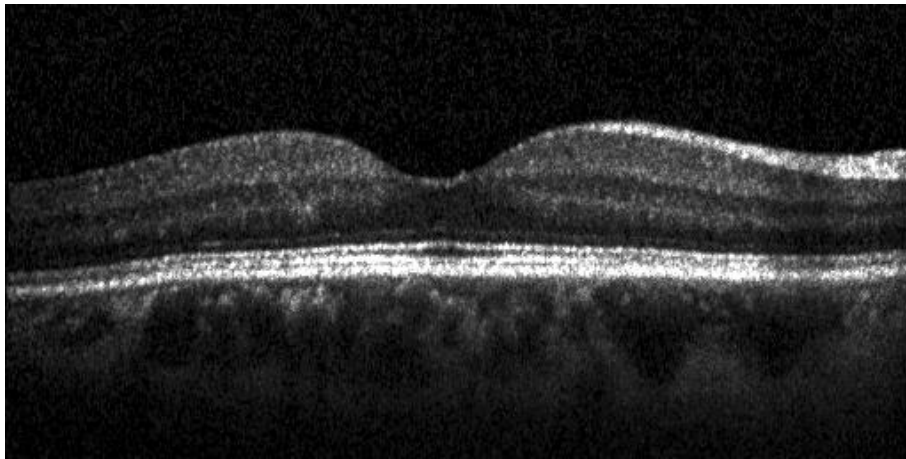


0_00076183_11.png

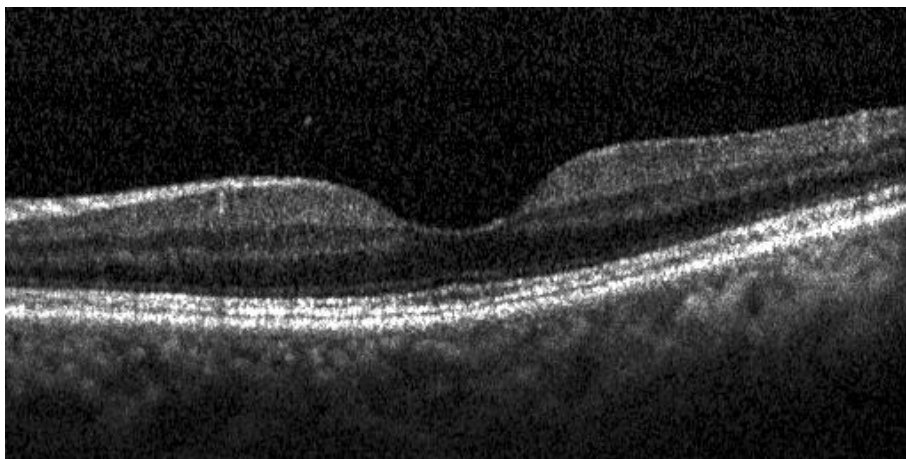
SimCLR Pre-train Model



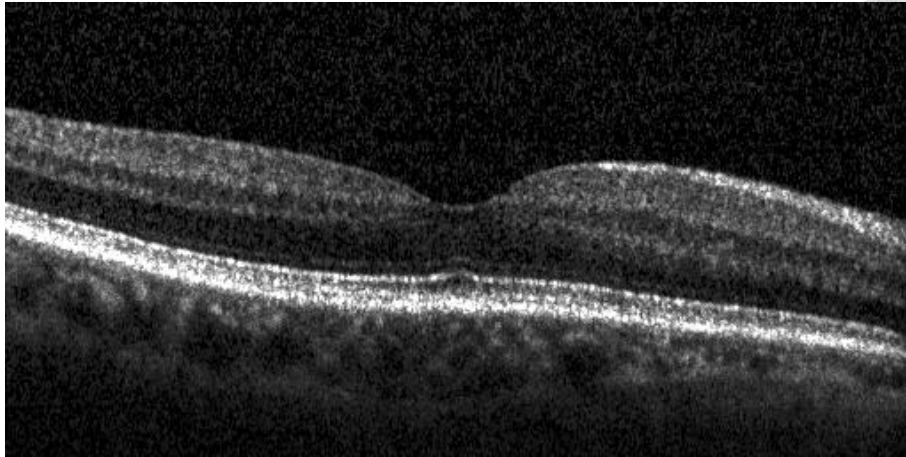
0_00078172_11.png



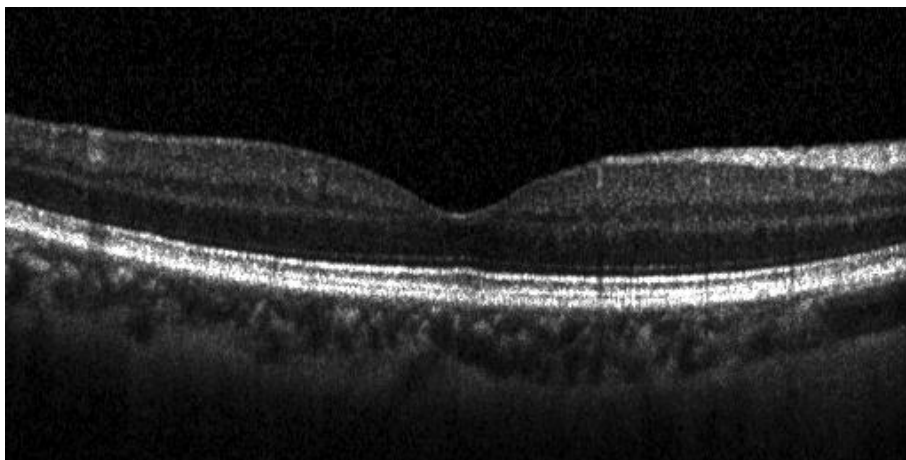
0_00061363_11.png



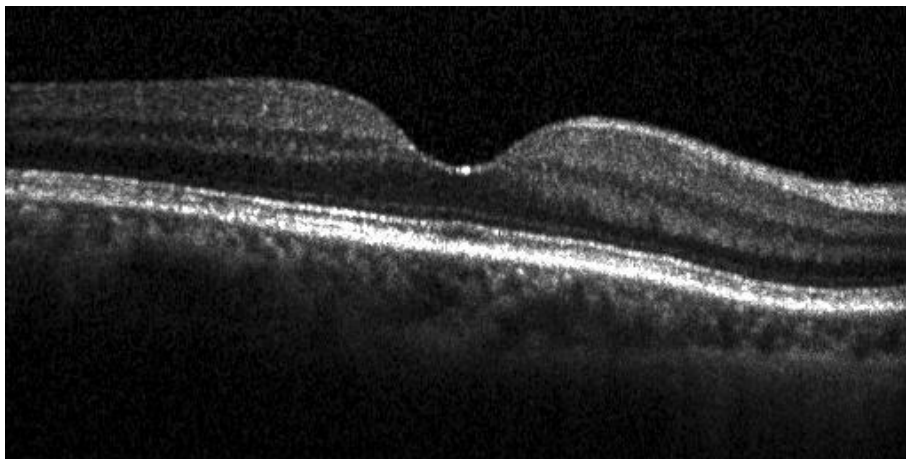
0_00085895_11.png



0_00092684_11.png

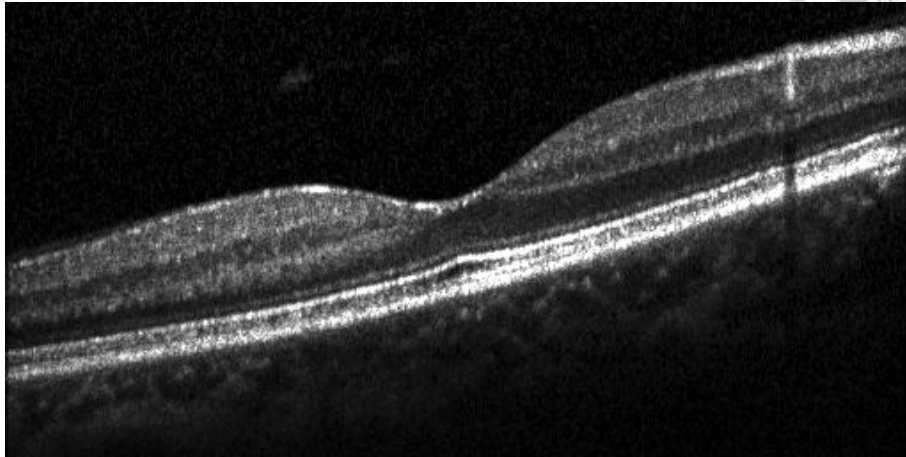
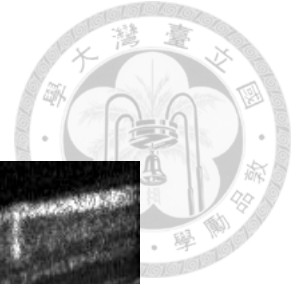


0_00110461_11.png

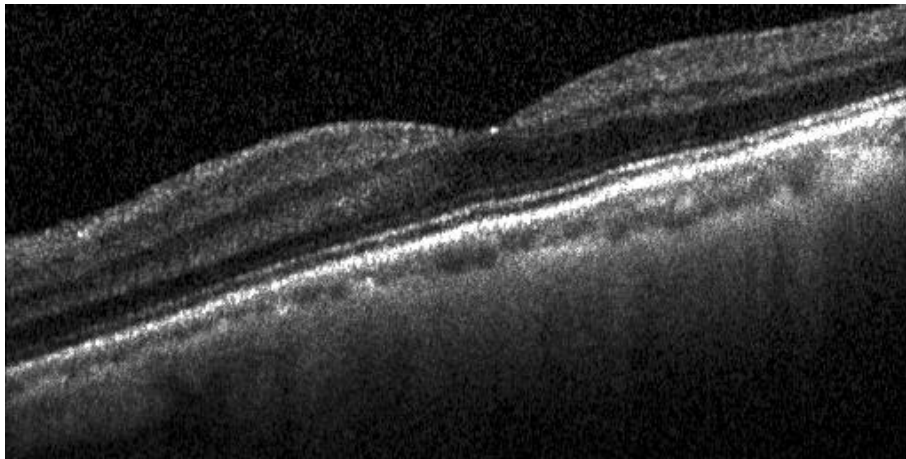


0_00122977_11.png

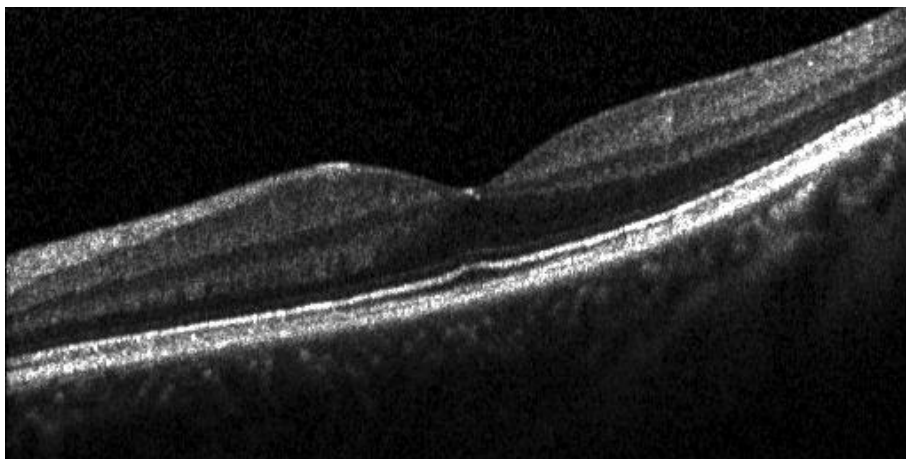
OCT Pre-train Model



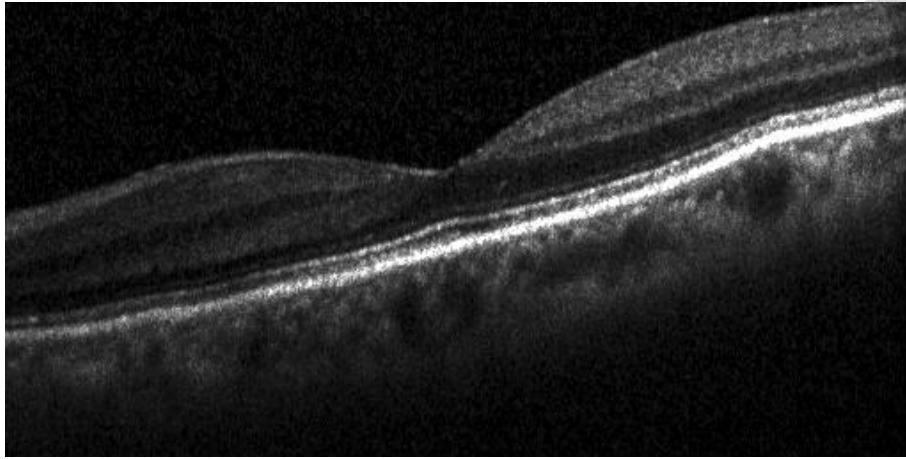
0_00035869_11.png



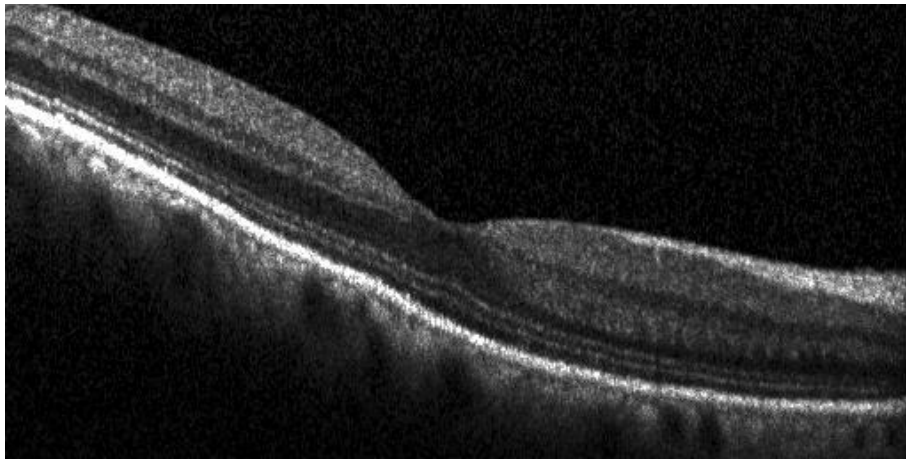
0_00040840_11.png



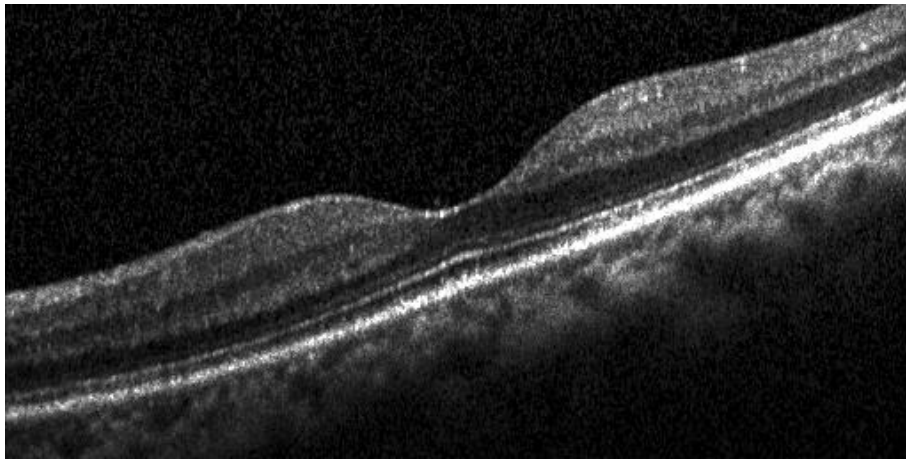
0_00049489_11.png



0_00071975_11.png



0_00096033_11.png



0_00099954_11.png