**中 國 文 化 大 學**

**資 訊 工 程 學 系**

**資 訊 系 統 專 題**

**糖尿病視網膜病變辨識系統**

**學 生：俞 華 鈴**

**李 馥 彤**

**林 永 泰**

**指導教授：林 世 崧**

**中 華 民 國 114年5月**

**糖尿病視網膜病變辨識系統**

**專題學生：俞華鈴、李馥彤、林永泰**

**指導教授：林世崧 博士**

**中國文化大學 資訊工程學系**

# 摘要

隨著現代社會飲食習慣與生活型態改變，糖尿病患者人數快速上升，導致因糖尿病引起的視網膜病變成為全球成人失明的主要原因之一。早期發現與治療是避免視力喪失的關鍵，特別是在醫療資源較匱乏的偏鄉地區，診斷延遲問題尤為嚴重。本專題旨在建構一套基於深度學習技術的糖尿病視網膜病變自動辨識系統，協助醫師提升診斷效率，並縮短偏鄉醫療落差。系統採用免散瞳眼底影像作為輸入，最終目標為提供一個可實際應用於初步篩檢的輔助診斷系統，降低病患視力喪失的風險，並減輕後續醫療負擔。

**關鍵詞**：深度學習、卷積神經網路、資料前處理與增強、影像辨識

指導教授\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_(簽名)

目錄

[摘要 I](#_Toc197996669)

[表 目 錄 IV](#_Toc197996670)

[圖 目 錄 V](#_Toc197996671)

[**第1章 研究動機與目的** 1](#_Toc197996672)

[1.1 研究動機 1](#_Toc197996673)

[1.2 研究目的 1](#_Toc197996674)

[1.3 研究流程 1](#_Toc197996675)

[**第2章 文獻探討** 3](#_Toc197996676)

[2.1 糖尿病 3](#_Toc197996677)

[2.1.1糖尿病併發症 3](#_Toc197996678)

[2.1.2糖尿病型視網膜病變 4](#_Toc197996679)

[2.2 糖尿病視網膜病變公開據集 6](#_Toc197996680)

[2.3影像處理 7](#_Toc197996681)

[2.3.1高斯濾波器 8](#_Toc197996682)

[2.3.2限制對比度自適應直方圖均衡化(CLAHE) 8](#_Toc197996683)

[2.3.3伽瑪校正 9](#_Toc197996684)

[2.4影像分類演算法 9](#_Toc197996685)

[2.4.1機器學習 10](#_Toc197996686)

[2.4.2深度學習 10](#_Toc197996687)

[2.4.3遷移式學習 11](#_Toc197996688)

[2.5卷積神經網路預訓練模型 12](#_Toc197996689)

[2.5.1 ResNet50 12](#_Toc197996690)

[2.5.2 VGGNet 14](#_Toc197996691)

[**第3章 研究內容、方法** 16](#_Toc197996692)

[3.1 研究內容 17](#_Toc197996693)

[**第4章 人力配置** 18](#_Toc197996694)

[**第5章 預期成果** 19](#_Toc197996695)

[**第6章 執行進度** 20](#_Toc197996696)

[**參考文獻** 21](#_Toc197996697)

[附錄A 開發工具介紹 23](#_Toc197996698)

[A.1 軟體開發工具介紹 23](#_Toc197996699)

[A.1.1 PyTorch 23](#_Toc197996700)

[A.1.2 Visual Studio Code 24](#_Toc197996701)

[A.1.3 TensorFlow 25](#_Toc197996702)

[A.1.4 Colab 25](#_Toc197996703)

[A.2 影像辨識電腦硬體 26](#_Toc197996704)

# 表 目 錄

[表 2.1 糖尿病視網膜病變臨床症狀 4](#_Toc197803828)

[表 2.2 DR分級 5](#_Toc197803829)

[表 2.3 糖尿病視網膜病變資料集 7](#_Toc197803830)

[表 2.4 糖尿病視網膜病變各等級圖片數量之分布 7](#_Toc197803831)

# 圖 目 錄

[圖 1.1 研究流程圖 2](#_Toc197803856)

[圖2.1 AI、ML、DL關係比較 10](#_Toc197803897)

[圖2.2 多層神經網路 11](#_Toc197803898)

[圖2.3 Layer Transfer 12](#_Toc197803899)

[圖2.4 ResNet50模型架構圖 13](#_Toc197803900)

[圖2.5 VGG16模型架構圖 15](#_Toc197803901)

**第1章 研究動機與目的**

## 1.1 研究動機

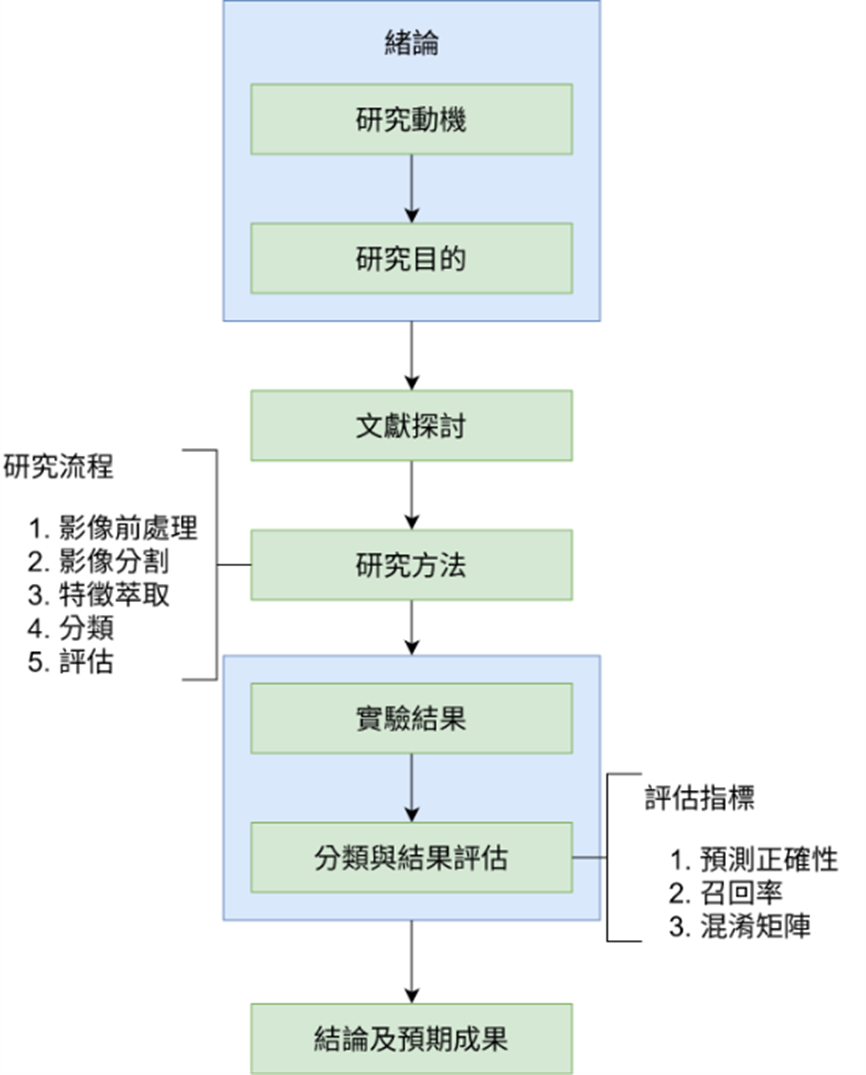
隨著全球糖尿病患者人數不斷攀升，糖尿病視網膜病變已成為成年人失明的主要原因之一。尤其在偏遠地區，受限於醫療資源不足與專業診斷人員匱乏，病患往往無法及時接受檢查與治療，導致視力不可逆損傷。

傳統眼底檢查需依賴專業醫師目視判讀，程序耗時且受限於設備與技術普及程度。隨著深度學習與醫療影像技術的進步，利用人工智慧協助糖尿病視網膜病變的早期偵測成為可行且迫切的需求。因此，本研究期望藉由科技手段縮短診斷流程，提升篩檢覆蓋率，減緩糖尿病引發的失明問題，改善偏鄉醫療可及性。

## 1.2 研究目的

本研究旨在開發一套基於深度學習技術的糖尿病視網膜病變自動辨識系統，使用眼底影像進行五級分級診斷，輔助醫療人員進行初步篩檢。透過卷積神經網路（CNN）模型訓練與優化，提高影像辨識的準確率與穩定性，並結合雲端平台，實現遠端即時診斷功能。最終目標為建立一個高效、準確且易於部署的醫療輔助系統，以提升早期診斷率，降低視力損失風險，並促進偏鄉地區醫療服務之普及與提升。

## 1.3 研究流程

本論文研究流程圖如圖1.1所示。首先介紹糖尿病視網膜病變的研究背景及相關議題，其中重點著重在於探討傳統人工檢測和自動化檢測的研究議題，找出糖尿病視網膜病變檢測需要解決及改善的問題。於文獻部分，將由現有的糖尿病視網膜病變自動化流程相關文獻整理，並探討過去學者所提出的技術，最後在整理的文獻範圍內篩選並明確研究主題。接著建立本研究架構，流程包括影像前處理、實驗模型評估、實驗模型參數調整。並開發網頁式糖尿病視網膜辨識系統，提供醫生迅速且準確地判斷視網膜病變程度，從而達到早期治療的目標。

**圖 1.1 研究流程圖**

**第2章 文獻探討**

## 2.1 糖尿病

當代社會，糖尿病不再是少數人的困擾，而是全球性的流行病。根據 J. W. Y. Yau 等人（2012）[1]的統計，全球已有約 9300 萬人患有糖尿病視網膜病變（DR），其中包括 1700 萬人罹患增殖性視網膜病變（PDR）、2100 萬人患有糖尿病黃斑部病變（DME），以及 2800 萬人正面臨威脅視力的視網膜病變（VTDR）。隨著糖尿病患者人數快速攀升，這些數字還在持續增加。

根據國際糖尿病聯盟（International Diabetes Federation, 2025）[2]的估計，2024 年全球約有 5.89 億名 20 至 79 歲的成年人患有糖尿病，佔該年齡層人口的 11.1%，也就是 每 9 人中就有 1 人。這樣的趨勢顯示，糖尿病正以驚人的速度成長，不僅是醫療問題，更是公共衛生的重大挑戰。

### 2.1.1糖尿病併發症

根據國際糖尿病聯盟（International Diabetes Federation, 2025）[2]的統計，到了2025年，全球糖尿病患者數量將突破3億人。糖尿病不僅造成血糖失控，更可能引發心血管、腎臟、神經系統及視覺系統等多重器官的併發症，從生活品質下降，到失明、洗腎甚至截肢，對患者個人、家庭乃至整個社會帶來沉重衝擊。

糖尿病是一種因血糖濃度過高、胰島素分泌不足或作用異常所引發的代謝性疾病。依據類型不同，糖尿病的併發症可分為急性與慢性兩大類[4]。急性併發症中最常見的是血糖控制不穩，特別是低血糖症，患者一旦出現症狀，需即時補充含糖食物以緩解危機。慢性併發症則影響更為深遠，涵蓋眼睛、腎臟、心血管以及神經系統等，長期發展下可能導致不可逆的器官損傷。

其中，糖尿病視網膜病變（Diabetic Retinopathy, DR）不僅涉及血管異常，如血管通透性增加、微血管阻塞及新生血管形成，也包含視網膜神經細胞的功能障礙與退行性變化，且神經病變可能早於血管病變出現[5]，上述糖尿病視網膜病變臨床症狀如表2.1所示。

**表 2.1 糖尿病視網膜病變臨床症狀**

|  |  |
| --- | --- |
| DR 嚴重等級 | DR 眼底影像特徵 |
| 無視網膜病變(No DR) | 無明顯視網膜病變 |
| 輕度非增殖性視網膜病變  (Mild NPDR) | 僅有微動脈瘤 |
| 中度非增殖性視網膜病變  (Moderate NPDR) | 微動脈瘤和出血較多，可能伴隨硬性滲出物 |
| 重度非增殖性視網膜病變  (Server NPDR) | 4-2-1法則，四個象限都有出血和微動脈瘤，或至少兩個象限有靜脈串珠狀改變，或至少一個象限有明顯的視網膜內微血管異常(IRMA) |
| 增值性視網膜病變  (PDR) | 視網膜新生血管形成、玻璃體或視網膜前出血、纖維增生 |

### 2.1.2糖尿病型視網膜病變

視網膜病變是成年人失明的主要原因之。根據近年研究，第1型糖尿病患者罹患視網膜病變的總發生率約為18%，年發生率近6%[6]；而第2型糖尿病患者的總發生率則為7.6%，年發生率約2.6%[7]。這些數字提醒我們，血糖控制、飲食管理、規律運動與良好的生活習慣，對延緩視網膜惡化至關重要。

糖尿病視網膜病變大致分為兩大類型，各等級眼底影像示意圖如下（見表2.2）：

非增殖性糖尿病視網膜病變（Non-Proliferative Diabetic Retinopathy, NPDR）：   
早期的微血管病變，包括微血管瘤、棉絮斑、硬性滲出物與點狀出血。依病變嚴重度分為輕度、中度及重度。

增殖性糖尿病視網膜病變（Proliferative Diabetic Retinopathy, PDR）：   
當視網膜缺氧惡化，異常新生血管開始滋生，脆弱的新血管容易破裂出血，進而引發玻璃體出血、視網膜剝離，嚴重者最終可能失明。

**表 2.2 DR分級**

|  |  |
| --- | --- |
| DR嚴重分級 | DR眼底影像 |
| 無視網膜病變(No DR) |  |
| 輕度非增值性視網膜病變  (Mild NPDR) |  |
| 中度非增值性視網膜病變  (Moderate NPDR) |  |
| 重度增值性視網膜病變  (Sever NPDR) |  |
| 增值性視網膜病變  (PDR) |  |

## 2.2 糖尿病視網膜病變公開據集

本節介紹本研究所使用糖尿病視網膜病變分級任務上資料集[8][9]。在分級實驗中，這些數據將被切分為訓練集、測試集及驗證集，測試及和驗證及通常用於評估系統效能，進一步用來與其他研究系統效能上進行比較。其中資料集依據疾病嚴重程度分為五類，本論文探討及研究表3中五等級分類之糖尿病視網膜病變資料集。

糖尿病視網膜病變資料集的資訊如表2.3，內容包含訓練及與測試集的張數、是否為多個專家標記等資訊。表2.4為各分級之資料集的各分類圖片分布數量，除了IDRiD資料集為固定的分辨率(4288\*4288)，APTOS 2019資料集則為不固定影像分辨率。APTOS 2019資料集的總圖片數量為5,590張，其中包含標記3,662張。

**表 2.3 糖尿病視網膜病變資料集**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **資料庫**​ | **訓練集(張)** | **測試集(張)** | **總數(張)**​ | **分類數量** | **是否標記** |
| IDRiD[10] | 413 | 103 | 513 | 5 | 是 |
| APTOS 2019[11] | 3662 | 1,928(因比賽結束未公開)​ | 5590 | 5 | 是 |

**表 2. 4 糖尿病視網膜病變各等級圖片數量之分布**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 資料庫​ | 訓練/測試集 | 分類 | | | | | 數量總計 |
| 0 | 1 | 2 | 3 | 4 |
| IDRiD | 訓練集 | 134 | 20 | 136 | 74 | 49 | 413 |
| 測試集 | 34 | 5 | 32 | 19 | 13 | 103 |
| APTOS 2019 | 訓練集 | 1,805 | 370 | 999 | 193 | 295 | 3,662​ |

## 2.3影像處理

影像處理是一種針對影像進行優化與分析的技術，目的是強化影像中的關鍵特徵，提升其辨識與應用價值。常見處理包括色彩增強、雜訊消除、特徵突出與資料擴增，藉此提升後續分析與判斷的準確性與穩定性。結合人工智慧後，影像處理技術被廣泛應用於人臉辨識、車牌辨識、醫學診斷、衛星監測與工業檢測等領域，成為智慧化應用不可或缺的基礎。

## 2.3.1高斯濾波器

高斯濾波器（Gaussian Filter）是一種常用於影像處理中的線性平滑濾波器，其主要目的是降低影像中的高頻雜訊，達到去雜與模糊（blurring）效果，並保留整體影像的輪廓與結構。此濾波器以高斯分佈函數（Gaussian distribution）作為權重基礎，其數學形式為：

其中，x 與 y 為相對於中心像素的空間位置，σ則為標準差，控制濾波器的模糊程度。標準差越大，濾波效果越明顯，平滑範圍越廣。高斯濾波器具有平滑連續邊緣、抑制隨機雜訊的能力，廣泛應用於影像前處理階段，例如邊緣偵測、特徵提取與醫學影像增強等任務。與均值濾波器（Mean Filter）相比，高斯濾波器能提供更自然的模糊效果，並減少邊緣失真，尤其適合應用於需要精細結構保留的影像處理場景。

## 2.3.2限制對比度自適應直方圖均衡化(CLAHE)

CLAHE（Contrast Limited Adaptive Histogram Equalization）是一種進階的影像對比增強技術，用於改善傳統直方圖均衡化（Histogram Equalization）在局部細節與雜訊控制上的限制。其核心概念為將影像分割為若干小區塊（tiles），針對每個區塊獨立進行直方圖均衡化，以提升局部對比，並避免過度增強造成的假影（artifacts），與 AHE（Adaptive Histogram Equalization）不同，CLAHE 引入對比限制機制[12]，即設定一個「clip limit」來抑制某些區塊中像素值過度集中的情形，進一步避免雜訊放大問題。增強後的區塊再透過雙線性插值（bilinear interpolation）整合，確保整體影像平滑過渡、不產生邊界不連續現象。CLAHE 在醫學影像（如眼底影像、X 光片）、衛星遙測影像、低光源環境下的影像強化等應用中，能有效改善影像清晰度與細節呈現，進而提升後續特徵偵測與分類任務的準確性。

## 2.3.3伽瑪校正

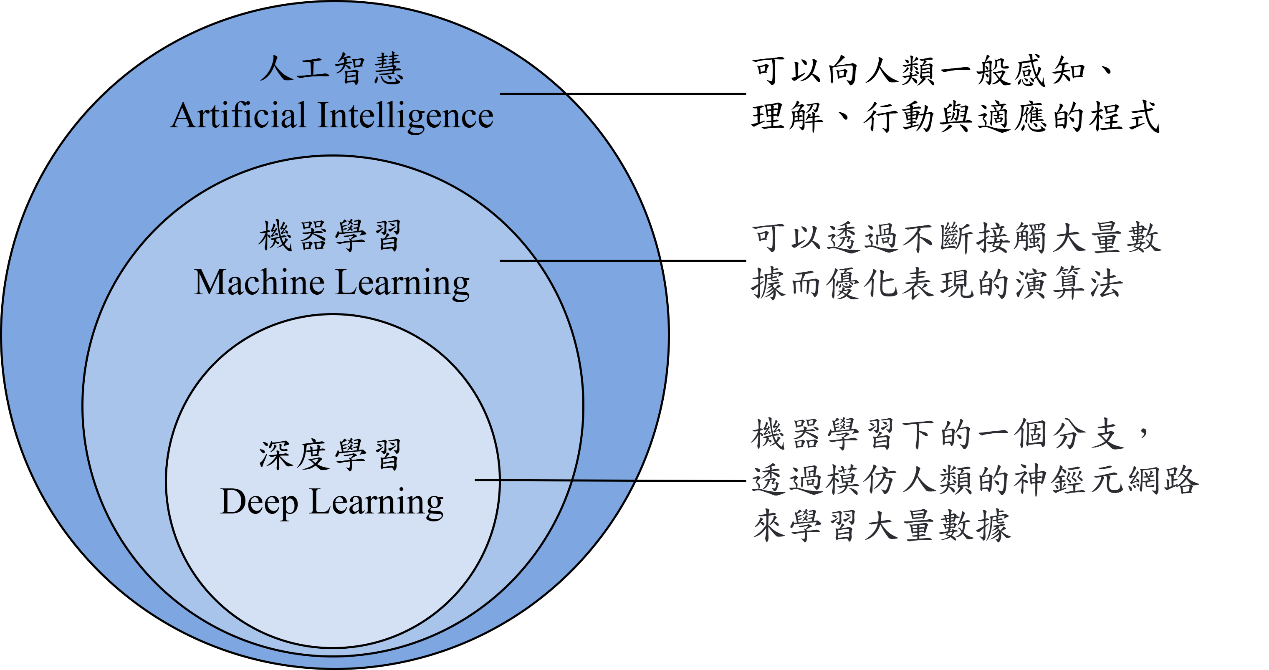
伽瑪校正（Gamma Correction）是一種非線性影像處理技術，主要應用於調整影像的亮度與對比，藉以符合人眼對光線強度的感知特性。人類視覺對亮度變化的感知是非線性的，對低亮度較為敏感，對高亮度則較為遲鈍。為了彌補這種非線性感知，伽瑪校正藉由非線性轉換公式，將原始像素值進行調整，使得影像在視覺上更自然、細節更清晰。其轉換公式為：

為原始像素值（通常已正規化為 [0,1] 範圍）。當 gamma <1 時，影像變亮；當gamma > 1 時，影像變暗。伽瑪校正在數位影像處理中廣泛應用，包括攝影曝光調整、顯示器影像校準、醫學影像增強及計算機視覺前處理等。透過適當的伽瑪值設定，可強調低亮度區域的細節，改善整體影像感知品質。

## 2.4影像分類演算法

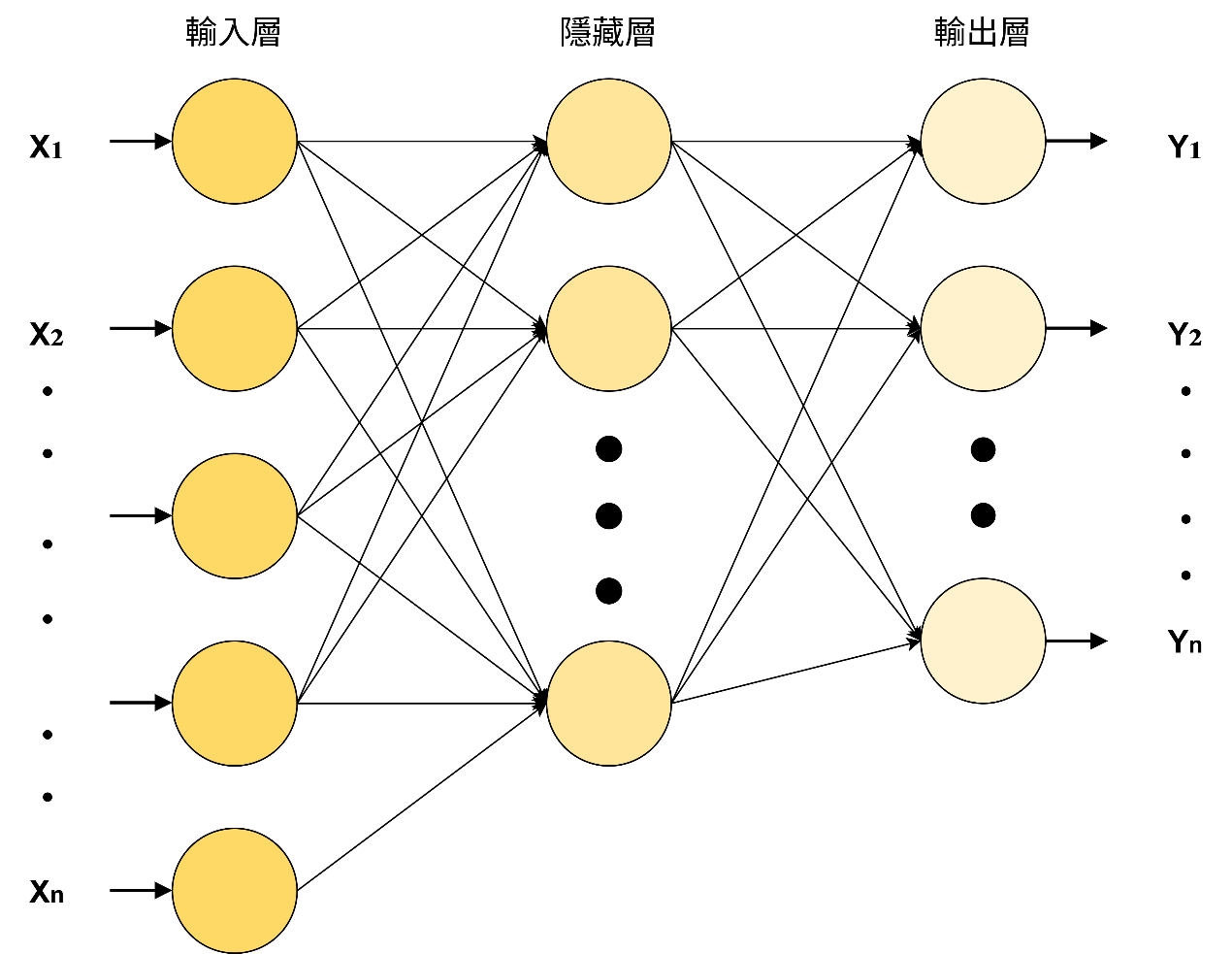
早期影像分類的做法，是人工一筆一筆資料標記特徵後丟進神經網路中進行分類，靠人眼與耐心堆疊。隨著 GPU 算力崛起，機器開始自己學：從大量資料中挖出模式，不管是傳統的機器學習(Machine Learning, ML)，還是只需少量線索就能深入理解的深度學習(Deep Learning, DL)，皆受到矚目，並廣泛的進行影像分類的應用，讓影像辨識變得更快、更準。

## 2.4.1機器學習

機器學習(Machine Learning, ML) 是人工智慧的核心技術之一，不再只是人給規則、機器照做，而是讓機器從數據中自己摸索、自己學。透過監督式學習、非監督式學習，甚至模擬人類試錯精神的增強式學習，機器開始理解語言、看懂影像、聽出聲音，也能預測市場、診斷疾病等。當人們給它標記過的資料，它學會用「監督式學習」來預測未來；當資料沒有標籤，它就用「非監督式學習」來自己歸類、找出結構。

**圖2.1 AI、ML、DL關係比較**

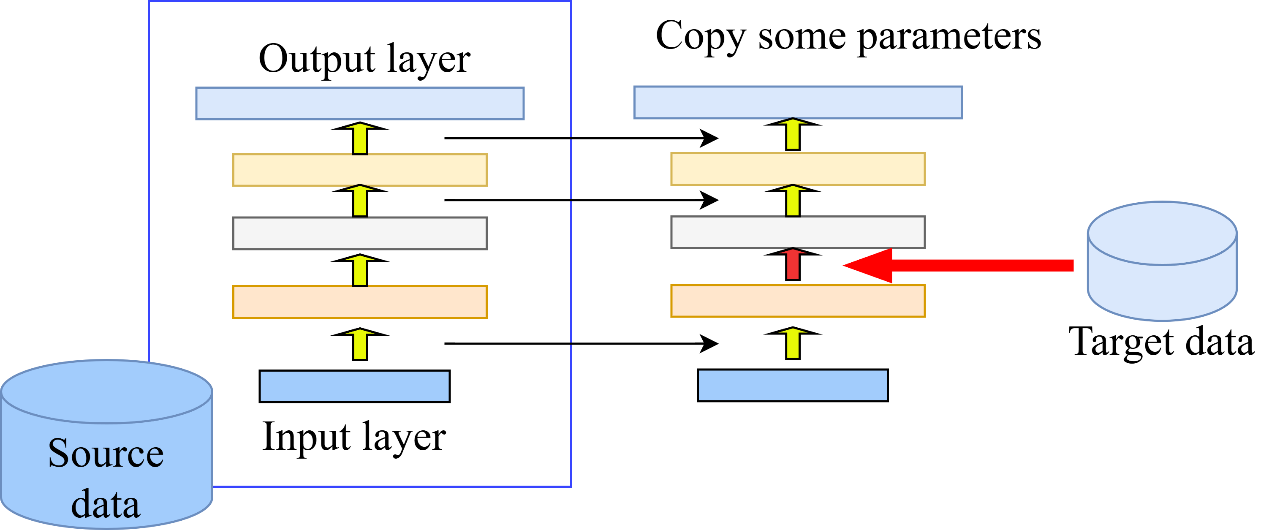
## 2.4.2深度學習

深度學習(Deep Learning, DL)是機器學習的一種方法，建立在人工神經網路的基礎上，透過多層結構模擬人腦思考過程。這些結構由輸入層、隱藏層與輸出層組成如圖3所示，每層通過特定的轉換單元提取數據特徵，並逐層傳遞與學習，讓模型在處理複雜任務與大量數據時更精準。卷積神經網路（CNN）是處理影像的核心架構之一，藉由自動提取特徵與反向傳播學習來完成分類與預測任務。隨著研究深入，架構如 ResNet、DenseNet 等透過增加網路深度與設計改良，已廣泛應用於圖像辨識、語音識別、語言翻譯等領域，展現高度準確與靈活性。

**圖2.2 多層神經網路**

## 2.4.3遷移式學習

遷移式學習（Transfer Learning）是一種機器學習技術，它的核心概念是：將一個任務上學到的知識，轉移到另一個相關任務上。不需要每次都從頭訓練模型，而是可以把已經在大量資料上訓練好的模型，像是 ImageNet 上的 ResNet、VGG、EfficientNet 等，拿來「微調（fine-tune）」或「凍結部分層」後，應用在新任務上，圖2.3為遷移式學習的基本流程，模型首先於來源資料上訓練，取得具代表性的特徵表達，將前面幾層保留，代換掉後面幾層之後重新訓練模型的參數，接著將部分網路層（如卷積層）之參數遷移至目標任務中。針對目標資料，可選擇僅訓練輸出層以降低過擬合風險，或在資料充足情況下對整體網路進行微調（fine-tuning），以進一步提升預測準確率。



**圖2.3 Layer Transfer**

## 2.5卷積神經網路預訓練模型

## 2.5.1 ResNet50

ResNet（Residual Network）由 Microsoft 研究員 Kaiming He 等人於 2015 年提出，並於同年在 ImageNet 大型視覺辨識挑戰（ILSVRC 2015）中榮獲分類任務冠軍[13]。該架構的核心創新在於引入「殘差學習（Residual Learning）」機制，有效解決了深層神經網路在層數增加時，可能導致訓練準確率反而下降的問題。相較於 VGGNet，ResNet 雖然層數更多，但因採用大量 1×1 卷積與參數共享策略，整體參數量與計算量反而更低。在訓練過程中，當網路已趨近最佳狀態時，若繼續加深卷積層，傳統模型容易出現退化現象（Degradation Problem）。ResNet 透過設計「殘差對映（Residual Mapping）」與「自身對映（Identity Mapping）」結構，使網路即使無法進一步學習新特徵時，也能透過跳接結構（skip connection）直接將輸入訊號傳遞至下一層。本研究選擇 ResNet 的經典ResNet50，作為分類糖尿病視網膜病變影像的基礎模型。ResNet-50 具備 50 層深度，結合殘差模組與全域平均池化（Global Average Pooling），在影像分類一張含有 文字, 螢幕擷取畫面, 字型, 數字 的圖片

AI 產生的內容可能不正確。任務中展現出優異的特徵擷取與泛化能力，其結構如圖 2.4所示。

**圖2. 4 ResNet50模型架構圖**

## 2.5.2 VGGNet

VGGNet 由牛津大學計算機視覺組（Visual Geometry Group, VGG）與 Google DeepMind 於 2014 年共同提出，並在ImageNet 大型視覺辨識挑戰賽（ILSVRC 2014）中榮獲分類任務第二名[12]。該模型在當時被視為極具深度的卷積神經網路架構之一，其最大特點為採用統一設計的卷積核與層級結構，顯著簡化了網路設計的複雜度。相較於早期的 AlexNet，VGGNet 採用多層 3×3 小卷積核取代單層 5×5 或 7×7 大卷積核的設計。這種策略不僅提升了非線性表達能力，還能在保持感受野的前提下減少參數量並增加深度。卷積層之間穿插使用 2×2 最大池化層進行空間下採樣，逐步提取並濃縮影像特徵。此一致性的架構設計使 VGGNet 容易實作與優化，並促使其成為後續許多深度學習架構設計的基礎模板。本研究選用VGG-16模型，作為糖尿病視網膜病變影像分類的基礎模型。如圖2.5所示VGG-16 包含 13 層卷積層、5 層最大池化層與 3 層全連接層，整體結構深度適中，具備優秀的特徵提取能力與穩定的分類表現，特別適合用於中小型醫學影像資料集的遷移學習應用。

一張含有 文字, 螢幕擷取畫面, 字型, 數字 的圖片

AI 產生的內容可能不正確。

**圖2. 5 VGG16模型架構圖**

**第3章 研究內容、方法**

一張含有 文字, 螢幕擷取畫面, 字型, 圖表 的圖片

AI 產生的內容可能不正確。本研究以Kaggle資料庫中APTOS 2019 Blindness Detection競賽所提供的眼底圖為資料來源，透過前處理統一影像規格後，將其輸入深度學習模型進行訓練。為提升實用性與可及性，特別開發一套基於網頁的糖尿病視網膜病變（Diabetic Retinopathy, DR）自動辨識系統，方便醫師上傳眼底圖並快速獲得診斷結果。DR 是糖尿病患者常見且嚴重的併發症，長期血糖控制不良會導致視網膜微血管病變，嚴重者甚至失明。傳統診斷仰賴專業醫師進行影像判讀，易受限於人力與設備，在偏鄉地區更常面臨診斷延遲問題。隨著人工智慧與深度學習技術的進展，本研究結合醫療影像處理與AI模型，自動辨識眼底圖中的病變等級，期望提升初步篩檢效率、減輕醫療人員負擔，並縮小偏鄉與都會區的醫療落差。圖3.1為本研究的糖尿病視網膜辨識系統架構圖。

**圖3. 1 糖尿病視網膜辨識系統架構圖**

## 3.1 研究內容

本研究的範圍涵蓋從資料蒐集、影像前處理、深度學習模型建構如圖3.2所示，到雲端部署與網頁前端整合，具體包括以下幾點：

**資料來源：**使用公開糖尿病視網膜病變影像資料集（ IDRiD、APTOS 2019），涵蓋不同等級之眼底影像。

**模型建立：**應用卷積神經網路（CNN）進行五級病變分類，包括無視網膜病變（No DR)、輕度非增殖性視網膜病變（Mild NPDR)、中度非增殖性視網膜病變（Moderate NPDR）、重度非增殖性視網膜病變（Server NPDR）及增值性視網膜病變（PDR）。

**系統架構：**將模型部署至雲端平台，並建置網頁介面供醫療人員或一般使用者上傳眼底影像獲取即時診斷結果。

一張含有 文字, 螢幕擷取畫面, 行, 字型 的圖片

AI 產生的內容可能不正確。**應用場景：**本系統特別針對偏鄉醫療場域設計，協助進行初步大規模篩檢，提升早期發現與治療機會。

**圖3.2系統架構圖**

**第4章 人力配置**

參與本研究之工作人員及預定之工作內容如表4.1所示。

**表4. 1參與本項研究之人員及工作內容**

|  |  |
| --- | --- |
| **姓名** | **工作內容** |
| 俞華鈴 | 資料蒐集、模型訓練與優化、  資料整理與報告統整 |
| 李馥彤 | 資料蒐集、模型訓練與實驗分析  、美術設計與報告編排 |
| 林永泰 | 資料蒐集、系統架構設計、  網頁前端開發、美術設計 |

**第5章 預期成果**

**第6章 執行進度**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 2月 | 3月 | 4月 | 5月 | 6月 | 7月 | 8月 | 9月 | 10月 | 11月 | 12月 | 1月 |
| 題目選定 |  |  |  |  |  |  |  |  |  |  |  |  |
| Kaggle  圖片資料集分析 |  |  |  |  |  |  |  |  |  |  |  |  |
| 資料雁整與處理 |  |  |  |  |  |  |  |  |  |  |  |  |
| 模型測試與調教 |  |  |  |  |  |  |  |  |  |  |  |  |
| 研究方法 |  |  |  |  |  |  |  |  |  |  |  |  |
| 論文撰寫 |  |  |  |  |  |  |  |  |  |  |  |  |

**參考文獻**

1. J. W. Y. Yau, S. L. Rogers, R. Kawasaki, E. L. Lamoureux, J. W. Kowalski, T. Bek, S.-J. Chen, J. M. Dekker, A. Fletcher, J. Grauslund, S. Haffner, R. F. Hamman, M. K. Ikram, T. Kayama, B. E. K. Klein, R. Klein, S. Krishnaiah, K. Mayurasakorn, J. P. O'Hare, T. J. Orchard, M. Porta, M. Rema, M. S. Roy, T. Sharma, J. Shaw, H. Taylor, J. M. Tielsch, R. Varma, J. J. Wang, N. Wang, S. West, L. Xu, M. Yasuda, X. Zhang, P. Mitchell, T. Y. Wong, and G. Meta-Analysis for Eye Disease Study(2012), “Global Prevalence And Major Risk Factors Of Diabetic Retinopathy,” Diabetes care, Vol. 35, No. 3, pp. 556-564, DOI: 10.2337/dc11-1909.
2. International Diabetes Federation. (2025). IDF Diabetes Atlas (11th ed., pp. 41–42). Brussels, Belgium: International Diabetes Federation. Retrieved from https://diabetesatlas.org/resources/idf-diabetes-atlas-2025/
3. 周碧瑟,董道興,李佳琳,莊紹源,林敬恆,楊南屏(2002), 「台灣地 區糖尿病流行病學」,台灣公共衛生雜誌,第二十一卷,第 83-96 頁,DOI: 10.6288/TJPH2002-21-02-01。
4. Alharbi, T. J., Alzahrani, F. M., Alghanmi, A. S., Alfatease, A. A., Alzahrani, S. H., & Alqarni, S. A. (2022). Acute and chronic diabetes-related complications among patients. Cureus, 14(10), e30376. https://doi.org/10.7759/cureus.30376
5. Tien, T. E., & Wong, Y. T. (2023). Diabetic Retinopathy: Looking forward to 2030. Frontiers in Endocrinology, 13, Article 1077669. DOI:10.3389/fendo.2022.1077669
6. Yau, J. W. Y., Lee, P., Wong, T. Y., Best, J., & Jenkins, A. (2023). Global prevalence and major risk factors of diabetic retinopathy. British Journal of Ophthalmology, 107(6), 742–748. doi:10.1136/bjophthalmol-2023-321123
7. Saaddine, J. B., Honeycutt, A., Narayan, K. M., Zhang, X., Klein, R., & Boyle, J. P. (2021). Projection of diabetic retinopathy and other major eye diseases among people with diabetes mellitus in the United States. JAMA Ophthalmology, 139(5), 546–552. doi:10.1001/jamaophthalmol.2021.1029
8. Srivastava, S., & Singh, A. (2021). Diabetic retinopathy detection using deep neural networks. In V. K. Gunjan, M. Zurada, & A. K. Sangaiah (Eds.), Proceedings of International Conference on Recent Advancement on Computer and Communication (pp. 697–705). Springer. https://doi.org/10.1007/978-981-15-5788-0\_64
9. Pratt, H., Coenen, F., Broadbent, D. M., Harding, S. P., & Zheng, Y. (2018). Convolutional neural networks for diabetic retinopathy. Data, 3(3), 25. https://doi.org/10.3390/data3030025
10. Indian Diabetic Retinopathy Image Dataset (IDRiD)

https://ieee-dataport.org/open-access/indian-diabetic-retinopathy-image-dataset-idrid

1. APTOS 2019 Blindness Detection – Kaggle:

https://www.kaggle.com/c/aptos2019-blindness-detection

1. Zuiderveld, K. (1994). Contrast limited adaptive histogram equalization. In P. Heckbert (Ed.), Graphics Gems IV (pp. 474–485). Academic Press. https://doi.org/10.1016/B978-0-12-336155-4.50061-6
2. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778. https://doi.org/10.1109/CVPR.2016.90
3. Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. International Conference on Learning Representations (ICLR). https://arxiv.org/abs/1409.1556

# 附錄A 開發工具介紹

## A.1 軟體開發工具介紹

### A.1.1 PyTorch

一張含有 圖形, 字型, 標誌, 平面設計 的圖片

AI 產生的內容可能不正確。PyTorch 是由 Facebook AI Research（FAIR）於 2016 年發布的開源深度學習框架，採用動態計算圖（Dynamic Computational Graph）結構，允許用戶在前向傳播階段隨時構建和修改網路，從而提高模型開發的靈活性和調試效率。其核心組件包括：自動微分引擎（Autograd），可追蹤張量運算並自動計算梯度；高效的張量（Tensor）操作，可在 CPU 與 NVIDIA GPU 之間無縫切換；以及豐富的生態庫，如 TorchVision（影像處理）、TorchText（文本處理）和 TorchAudio（聲音處理），以簡化常見任務的數據加載與預處理。此外，PyTorch 支援分散式訓練（torch.distributed），可在多 GPU 或多節點環境中擴展訓練；並通過 TorchScript 將模型序列化為靜態圖，方便在 Python 以外的環境進行部署。結合社群活躍的開發和廣泛的教學資源，PyTorch 已成為研究人員與工程師進行快速原型驗證和生產環境部署的首選深度學習平台，如圖A.1。

**圖 A.1 PyTorch**

### A.1.2 Visual Studio Code

Visual Studio Code（簡稱 VS Code）是微軟於 2015 年推出的一款免費、開源且跨平台的輕量級程式編輯器，基於 Electron 框架構建，兼具效能與可擴展性。VS Code 預設內建智慧提示、語法高亮、程式片段（Snippets）、整合終端機與強大除錯功能，並透過 Language Server Protocol（LSP）提供多種程式語言的深度語意分析與自動補全。

使用者可在 Extensions Marketplace 中安裝數千款擴充套件，涵蓋 Python、JavaScript、TypeScript、C#、Go、Rust、Docker、Kubernetes、Markdown 等領域，同時支援 ESLint、Prettier、GitLens、Live Server 等開發與協作工具。VS Code 內建 Git 版控整合，可直接在側邊欄檢視檔案差異、執行提交與分支管理；Live Share 則讓團隊成員可即時共用編輯與除錯會話。

此外，VS Code 提供 Settings Sync 功能，能跨裝置同步使用者的設定、主題、鍵盤快速鍵與已安裝擴充；其 JSON 格式的設定檔結構清晰，易於版本控制與團隊共享。整合式工作區（Workspace）管理與多根程式碼視窗支援，讓大型專案開發與多任務切換變得流暢。憑藉其直觀介面、高度客製化與豐富生態系，VS Code 已成為全端開發者、資料科學家與 DevOps 工程師的首選編輯器，適用於原型開發、系統維護、測試與自動化部署等多種場景，顯著提升開發效率與協作體驗。

一張含有 螢幕擷取畫面, 圖形, 符號, 電子藍 的圖片

AI 產生的內容可能不正確。

**圖 A.2 Vs code**

### A.1.3 TensorFlow

TensorFlow 是 Google Brain 團隊於 2015 年開源推出的機器學習平台，採用計算圖機制與優化執行時引擎，可高效執行張量運算。它整合 Keras 高階 API，提供模型定義、訓練與評估一體化接口，同時支援 Eager Execution 增強開發靈活性。借助 tf.data 可構建高效數據管道，結合混合精度和 tf.distribute 分散式策略，實現多 GPU 與 TPU 集群擴展訓練。內建 TensorBoard 用於可視化指標、網絡結構及超參數調優；訓練完成後，可導出 SavedModel，並透過 TensorFlow Lite、TensorFlow.js 與 TensorFlow Serving 在行動端、瀏覽器與服務端實現一站式部署。豐富生態與跨平臺支持，使其成為從原型開發到大規模生產部署的理想選擇，如圖A.3。

一張含有 字型, 標誌, 圖形, 符號 的圖片

自動產生的描述

**圖 A.3 TensorFlow**

### A.1.4 Colab

Colab，全名Google Colaboratory，是由 Google 提供的一款免費且免安裝配置的雲端 Python 開發環境，專為數據科學和機器學習任務而設計。它基於 Jupyter Notebook，允許用戶在瀏覽器中直接撰寫並執行 Python 代碼，並預裝了 NumPy、Pandas、Matplotlib、Scikit-learn、TensorFlow、PyTorch 等主流科學計算與深度學習庫。使用者只需擁有 Google 帳戶，即可一鍵啟動運算環境，並在「執行階段設定」中選擇 NVIDIA GPU 或 Google TPU 加速器，以滿足各種高性能訓練需求。Colab 無縫整合了 Google Drive，讓用戶能方便地讀寫和備份資料集，並支援與 GitHub 的協同操作，以便於版本控制與團隊協作。此外，多人可同時在線編輯同一份 Notebook，並透過分享連結或導出為 .ipynb、HTML、PDF 等格式，輕鬆分發研究成果或教學內容。由於免除了本地環境配置的複雜性並提供了強大的雲端計算資源，Colab 成為初學者快速入門以及專業人士進行原型驗證和模型部署的首選平臺，如圖A.4。

## A.2 影像辨識電腦硬體

**表A.1影像辨識電腦1**

|  |  |
| --- | --- |
| 作業系統 | Windows 11 家用版 |
| 處理器 | Intel(R) Core(TM) i7-12700H |
| 晶片組型號 | Intel HM670 PCH |
| 標準記憶體 | 16GB |
| 記憶體 | DDR5-4800 |
| 外部顯卡 | NVIDIA GeForce RTX 4050 laptop GPU |