

元智大學資管系

第三十屆學術類畢業專題頂石課程(二)

期末報告

應用深度學習於精神疾病腦部 MRI 影像分析

1111617 鄭又瑄、1111629 劉喬恩、1111638 郭喬恩

實習公司：郭文嘉教授

工作代號：ZD1

指導教授：郭文嘉教授

中華民國 114 年 11 月

November, 2025

摘要

根據世界衛生組織 (WHO) 統計，全球每八人中就有一人患有精神疾病，但高達七成的患者未能獲得適當治療[1]。傳統的精神疾病診斷主要依賴臨床症狀評估與主觀判斷，容易受到醫生經驗的影響，缺乏客觀的影像學分析支持。因此，開發自動化、客觀且高準確率的精神疾病診斷輔助工具，對於提升診斷效率與準確性具有重要價值。

本研究旨在透過深度學習技術 (3D EfficientNet) 進行精神疾病 MRI 影像的分析，並探討腦影像特徵與不同精神疾病 (如思覺失調症、躁鬱症與重鬱症) 之間的關聯。我們將採用台灣健康老化與精神疾病腦影像資料庫 (TAMI)，分析 1,265 位受試者 (587 名健康者與 678 名患者) 的 MRI 影像數據，透過 3D EfficientNet 模型學習腦部結構異常區域，並驗證其在精神疾病診斷輔助上的可行性。

最終，本研究希望能提供一種客觀的影像學輔助分析方法，以增強精神疾病的早期篩檢能力，並為智慧醫療與影像診斷技術的發展提供參考。

目錄

摘要.....	1
目錄.....	2
Chapter 1 研究主題.....	3
Chapter 2 緒論.....	4
2.1 研究背景.....	4
2.2 研究動機及問題.....	5
Chapter 3 文獻回顧與探討.....	6
Chapter 4 研究方法及步驟.....	10
4.1 資料集.....	10
4.1.1 資料集概述.....	10
4.1.2 資料集分析.....	11
4.2 001.mgz 與 brain.mgz.....	13
4.3 3D EfficientNet 模型架構.....	14
4.4 CBAM.....	16
4.5 Grad-CAM.....	16
4.6 K-Fold Cross Validation (交叉驗證).....	17
4.7 Learning Rate 設定與學習率調度器 (Scheduler).....	18
Chapter 5 研究成果.....	18
5.1 健康與精神疾病 CBAM 二元分類.....	18
5.2 年齡預測.....	20
5.3 實驗流程.....	21
Chapter 6 實驗結果分析.....	22
Chapter 7 結論.....	23
文獻參考.....	24
附錄 A. 專題工作內容.....	26
附錄 B. 專題心得與建議.....	28

Chapter 1 研究主題

根據世界衛生組織（WHO）說明，精神疾病的診斷多依賴臨床對症狀的觀察與功能評估，加上診斷資源不足與病識感低落等問題，導致許多患者無法及時獲得正確診斷與治療[2]。隨著神經影像學與深度學習技術的進步，許多研究開始關注是否能從腦部結構影像中找出具有鑑別力的生物標記（biomarkers），以輔助精神疾病的早期判別與類型分化。

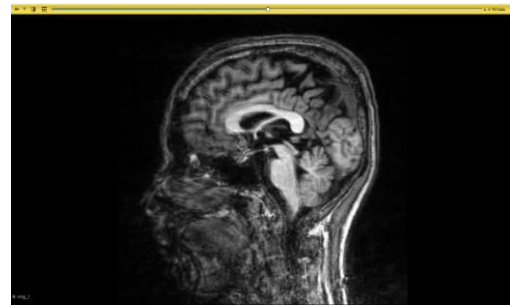
本研究旨在結合腦部 MRI 影像與 3D EfficientNet 深度模型，探討不同精神疾病——思覺失調症、躁鬱症與重鬱症——與腦部異常結構之關聯性，進而建立一套能有效分割病灶區域的影像辨識系統。我們將以台灣健康老化與精神疾病影像資料庫（TAMI）為基礎，進行多病類腦影像分析，期望提出具臨床潛力的診斷輔助工具。透過此研究，我們希望提供一種結合深度學習與醫學影像的可行解法，補足臨床現行診斷工具的不足，並為未來更客觀及標準化的精神疾病診斷提供重要參考與技術基礎。

預期研究成果：

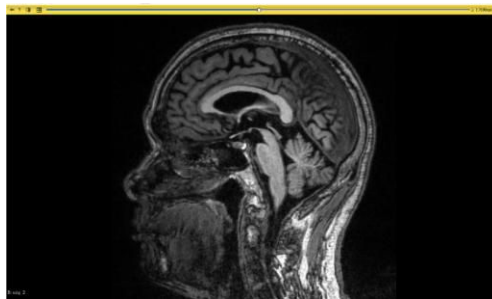
1. 建立基於 3D EfficientNet 針對精神疾病 MRI 影像分析模型，包括客觀分析流程與量化方法。
2. 為智慧醫療與影像診斷提供技術參考，推動醫療結合人工智慧在精神疾病診斷領域的應用。
3. 初步驗證深度學習應用於精神疾病影像分析的可行性，比較不同疾病組與健康對照組的差異。
4. 開發簡易的影像可視化工具，幫助直觀理解模型預測的異常區域分佈情況。



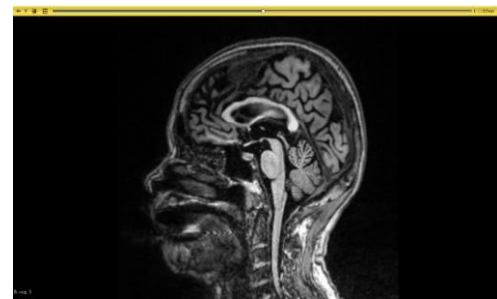
圖一、健康 001.mgz 樣本



圖二、思覺失調症 001.mgz 樣本



圖三、躁鬱症 001.mgz 樣本



圖四、重鬱症 001.mgz 樣本

Chapter 2 緒論

2.1 研究背景

精神疾病是全球公共衛生的重要挑戰之一，其複雜的病理機制與多樣化的臨床表現，使得診斷與治療需仰賴多面向的評估。當前，精神疾病的診斷主要依據國際通用的診斷標準（如 DSM-5 或 ICD-11），結合臨床醫師的專業判斷與患者的主觀症狀描述。此方式經過長期驗證，具有穩健性與臨床實用性，但診斷可能因主觀判斷而存在變異性，且 DSM-5 在美國使用較廣，ICD-11 則更普於全球應用[3]。近年來，隨著醫學影像技術的進步，研究發現許多精神疾病（如思覺失調症、重鬱症、躁鬱症等）可能伴隨腦部結構或功能的變化，例如思覺失調症的腦室擴大或重鬱症的海馬體積縮小。這些潛在的生物

標記 (biomarker) 雖具研究價值，但因其異質性高且缺乏特異性，目前尚未能作為常規診斷的可靠依據，僅在研究中展現未來潛力。

近年來，深度學習技術為醫學影像分析帶來突破，包括 3D U-Net、DeepMets 等模型能自動化分割腦部結構，高效提取細微形態特徵，大幅提升分析效率與可重現性[4]。然而，精神疾病的高度異質性與腦部異常的細微性，使模型難以辨識跨疾病的共享特徵，且需要大規模、標準化的影像數據以提升泛化能力，這仍是當前研究的核心挑戰。在此背景下，結合深度學習與腦影像分析，開發輔助診斷工具，旨在提供客觀的影像學證據，與 DSM-5 或 ICD-11 的臨床評估互補，而非取代傳統診斷方式，進而提升精神疾病診斷的精準度與客觀性。

2.2 研究動機及問題

2024 年震驚社會的臺北保母虐童致死案，一名一歲多的男童被寄養保母及其妹妹虐待長達 114 天，在這期間他被餵食餿水、惡意謾罵和暴力對待，最終不幸死亡。全身遍布多達 42 處的傷勢，包含多處瘀傷、牙齒斷裂和生殖器裂傷等，甚至患有憂鬱症和創傷後壓力症等心理創傷[5]。後來，保母開始宣稱自己長期有幻聽，可能患有思覺失調症，經專業醫生鑑定不符合精神病理診斷[5]。除了這個案件以外，近年來多個社會案件的被告都想透過精神疾病來減輕或免除刑責，但精神疾病的診斷判斷不易。

傳統的精神疾病診斷，大多透過醫生問診來釐清疾病史、家族史和篩檢量表等，並根據 DSM-5 診斷手冊或其他相關標準來判斷，缺乏客觀的影像學分析支持。除了病人未清楚表達症狀，也常因為發病時間不同，容易導致精神疾病的誤判，像是思覺失調症的早期症狀有時會誤判為躁鬱症或重鬱症，躁鬱症也會誤判為重鬱症，影響治療效果[6][7]。

本研究計畫預計使用 3D EfficientNet 深度學習技術，進行精神疾病 MRI 影像的分析，探討腦影像特徵與不同精神疾病之間的關聯，分別為思覺失調症 (Schizophrenia)、躁鬱症 (Bipolar Disorder，又稱雙向情感障礙) 和重鬱症 (Major Depressive Disorder)。協助醫生以更客觀和有效的方式精神疾病判別，及早發現真正罹患思覺失調症、躁鬱症與重鬱症的人，讓他們早日接受相關治療和恢復健康，未來也可以提供給精神疾病領域作為研究參考依據。

Chapter 3 文獻回顧與探討

本研究由 Albert C. Yang 等人於 2013 年發表於《Neurobiology of Aging》的《Complexity of spontaneous BOLD activity in default mode network is correlated with cognitive function in normal male elderly: a multiscale entropy analysis》[8]。研究聚焦於正常老年男性大腦預設模式網絡 (Default Mode Network, DMN) 中自發性血氧依賴訊號 (BOLD) 的時間複雜性，探討其與認知功能的關聯。研究採用多尺度熵 (Multiscale Entropy, MSE) 分析方法，首次應用於靜息態 BOLD 訊號，評估 DMN 活動的複雜性，並檢驗其與認知表現 (如記憶、注意力及執行功能) 的相關性。研究旨在探索隨著老化，大腦靜息態 BOLD 訊號的時間動態特性是否發生變化，並揭示其與認知能力的潛在關聯，為神經退化性疾病 (如阿茲海默症) 的早期診斷提供理論基礎。研究設定三個主要目標：(1) 探索 MSE 分析的適合參數與時間尺度；(2) 比較年輕與年長族群間 BOLD 訊號複雜度的差異；(3) 探討特定腦區的複雜度與認知功能的關聯。研究強調，靜息態 BOLD 訊號並非隨機雜訊，而是具有組織性與結構性，類似於心跳、步態等生理訊號隨老化表現出的複雜度下降趨勢。

研究招募 99 名正常老年男性受試者(平均年齡約 70 歲,無神經或精神疾病史),並納入 56 名年輕對照組進行比較。所有受試者接受功能性磁共振造影(fMRI)掃描,使用 T1 加權影像記錄靜息狀態下 DMN 的 BOLD 訊號(取樣率 0.4 Hz,TR=2.5 秒,時序資料點 195 點)。影像數據經過預處理,包括切割、對齊與校正,隨後針對自動化解剖標記(Automated Anatomical Labeling, AAL)定義的腦區,計算每個區域內體素的 BOLD 時間序列平均值,生成代表性時間序列。研究採用多尺度熵(MSE)分析,計算 BOLD 訊號在不同時間尺度(1 至 5 級)的樣本熵,作為神經活動複雜性的指標。MSE 參數選擇通過學生 t 檢驗優化,比較低認知分數與高認知分數受試者間各 AAL 區域的樣本熵差異,確定能夠有效區分組別的參數組合。認知功能評估包括迷你精神狀態檢查(MMSE)及其他認知量表,涵蓋記憶、注意力與執行功能。數據分析使用皮爾森相關分析檢驗 MSE 值與認知分數的關係,並以 Box-Cox 變換校正認知分數的偏態分佈,確保正態性;多元回歸模型控制年齡、教育程度等混淆變量。此外,研究比較年輕組與老年組的 MSE 差異,並分析 MSE 與變異性指標(標準差,SD)的相關性,以區分不同神經訊號特徵。

研究結果顯示,DMN 中 BOLD 訊號的 MSE 值與認知功能呈顯著正相關,特別是在短期記憶、執行功能與注意力方面,相關腦區包括海馬、旁海馬區、前額葉、尾核、丘腦與顳葉,其中以後扣帶皮質的變化最為顯著。MSE 值較高的受試者在認知測試中表現較佳,提示較高的神經活動複雜性可能反映更好的大腦功能儲備。與年輕組相比,年長組在多個腦區的 BOLD 複雜度顯著下降,尤以 DMN 內部最為明顯,這與老化導致 DMN 功能連結減弱的既有研究一致,支持「大腦網路整合能力隨年齡減弱」的假說。研究進一步指出,BOLD 訊號的複雜度下降可能與認知退化相關,與阿茲海默症等神經退化性疾病的發現相呼應。MSE 方法的應用展示了其量化神經訊號動態的潛力,有助於區分健康老化與病理性老化。然而,研究存在若干限制:(1) fMRI 時序資料僅 195 點,可能影響 MSE 計算的穩定性;(2) 年長組可能存在腦萎縮,影響訊號品質;(3) 取樣率較低(TR=2.5 秒),呼吸與心跳等生理雜訊可能干擾結果;(4) 僅納入男性受

試者，限制結果的普遍性；(5) 橫斷式設計無法推論長期變化趨勢。此外，MSE 與變異性指標 (SD) 的相關性不高，甚至在部分腦區呈反向趨勢，顯示兩者反映不同層面的神經特徵。未來研究可整合結構性影像 (如 DTI) 或功能連結分析，進一步闡明老化對大腦功能的影響機制。

研究表明，正常老年男性 DMN 中自發性 BOLD 訊號的時間複雜性與認知功能密切相關，特別是在記憶與執行功能方面，多尺度熵 (MSE) 分析是一種有效工具，可量化神經活動的動態特性。研究支持 DMN 在認知功能中的關鍵角色，並揭示 BOLD 訊號複雜度下降可能是正常老化的生理特徵，與認知退化及神經退化性疾病相關。這些發現為發展認知老化的神經生物標記提供了重要方向。後續研究應擴大樣本範圍，納入女性受試者，採用縱向設計追蹤複雜性變化的長期效應，並結合多模態影像 (如 DTI 與功能連結分析) 深入探索老化機制，以提升對健康與病理性老化的理解並促進早期診斷與干預策略的發展。

本研究由 J.D. Zhu 等人於 2023 年發表的《Investigating brain aging trajectory deviations in different brain regions of individuals with schizophrenia using multimodal magnetic resonance imaging and brain-age prediction: a multicenter study》[9]，旨在探討精神分裂症患者大腦老化軌跡的偏差，特別聚焦不同腦區的結構與功能變化。研究採用多模態磁共振造影 (MRI)，包括結構影像 (T1 加權，灰質 GM)、功能連結 (resting-state fMRI, FC) 與擴散張量影像 (DTI，分數各向異性 FA)，並利用 230 名健康對照者的數據訓練大腦年齡預測模型，作為比較精神分裂症患者與健康對照者腦齡差異的基準。研究發現，精神分裂症患者的的大腦老化速度顯著加快，尤其在額葉、顳葉、島葉等灰質區域，以及胼胝體、穹窿、小腦腳等白質束中，顯示出明顯的區域特異性老化標記。然而，功能連結圖未顯示顯著的加速老化，提示功能變化可能遵循不同機制。研究提出的腦年齡差距 (Brain Age Gap, BAG) 可作為精神分裂症疾病進展或治療監測的潛在生物標記，有助於理解其神經病理機制。此外，附加文件中提到的 TAMI (Taiwan Aging and

Mental Illness) 數據集，包含 194 名精神分裂症患者與 100 名健康對照者，進一步支持了本研究的多中心驗證方法。

研究採用多模態 MRI 數據結合機器學習，建構區域性腦齡預測系統，分析精神分裂症患者的腦齡差異，流程分為影像處理、模型建立與腦齡差分析三階段。在影像處理階段，整合 T1 加權（灰質）、靜息態 fMRI（功能連結）與 DTI（白質 FA）數據。結構影像經 DPABI 工具處理，依據 AAL 模板分割為 90 個灰質腦區；功能影像經 DPABI 與 MATLAB 進行時間序列處理與空間正規化，亦分割為 90 個功能區；白質纖維則使用 FSL 工具，依 JHU-ICBM atlas 分割為 48 條主要纖維路徑，共建立 228 個獨立腦齡預測模型。在模型建立階段，230 名健康對照者的數據用於訓練。特徵選擇採用皮爾森相關係數（ r ），針對各腦區隨機抽樣計算 100 次，篩選與年齡高度相關的 voxel（至少 10 次入選）。使用 Gaussian Process Regression（GPR）建立模型，通過五折交叉驗證評估效能，指標包括平均絕對誤差（MAE）與預測年齡相關係數。在腦齡差分析階段，模型應用於 TAMI（194 名精神分裂症患者、100 名健康對照者）與 BT（各 50 名患者與對照者）兩個隊列，計算腦齡差距（BAG）。統計分析使用獨立樣本 t 檢定與卡方檢定比較人口學與臨床變項差異，並以共變數分析（ANCOVA）比較 BAG，控制年齡、性別、MMSE 分數與教育年數，採用 FDR 校正（ $p < 0.05$ ）控制多重比較，效果量以部分 η^2 表示。BrainNet Viewer 用於結果視覺化。進一步以多元回歸分析檢驗 BAG 與臨床因子（如 PANSS 分數、病程、BMI）及抗精神病藥物劑量（CPZ 當量）的關聯性。

本研究揭示精神分裂症患者的腦老化軌跡顯著偏離正常老化，尤其在結構性指標上表現突出。在 90 個灰質腦區中，TAMI 隊列有 71 個區域、BT 隊列有 66 個區域的腦年齡差距（BAG）顯著大於健康對照者，額葉、顳葉與島葉為老化最嚴重的共同區域。在 48 個白質束模型中，TAMI 與 BT 隊列分別有 15 與 33 個白質束顯示老化偏差，包括胼胝體、穹窿、小腦中腳、左大腦腳、右矢狀束、左視丘腦後輻射、右穹窿/紋狀體、腦穹窿的身體與柱部及雙側前肢內囊等。此外，精神分裂症患者的發病年齡與四

個白質束的 BAG 呈負相關，提示疾病早期可能加速白質老化。本研究提出的區域特異性老化標記與腦年齡差距為精神分裂症的神經病理研究提供了重要線索，並作為潛在的疾病進展與治療監測指標。未來需通過縱向研究、多模態數據整合及更全面的臨床變量分析，進一步驗證與完善這些發現，以促進精神分裂症的早期診斷與個性化治療。

Chapter 4 研究方法及步驟

4.1 資料集

本研究採用臺灣老化與精神疾病序列 (Taiwan Aging and Mental Illness Cohort, TAMI) 的腦部 MRI 影像資料集[10]，該資料集由臺北榮總與陽明交通大學合作建置，主要收錄高齡精神疾病患者與正常對照組的結構性磁共振造影 (sMRI) 數據。資料格式以 .mgz (FreeSurfer 輸出格式) 為主，包含 3D 腦部影像，並搭配 Excel 表格記錄詳細標註資訊，如年齡、性別、臨床診斷，以及各腦區的體積、灰質/白質強度等量化數據。

4.1.1 資料集概述

在資料預處理階段，主要使用 001.mgz 和 brain.mgz，為確保模型輸入一致性，所有影像皆維持原始體素矩陣大小 $192 \times 256 \times 256$ 。

資料集劃分採用分層 20 折交叉驗證 (Stratified K=20)，輪流用 1 份來測試，其餘 19 份再切分為訓練集與驗證集 (80/20)，以解決樣本數不足的問題，同時確保拆分中類別比例一致並提升評估穩定性。

4.1.2 資料集分析

資料集總共分為一個健康組（443 筆資料）和三個精神疾病的資料集，包括思覺失調症（255 筆資料）、重鬱症（111 筆資料）和躁鬱症（139 筆資料），全部總計 948 筆資料。其中，所有的資料集都擁有部分共同欄位，如表一的資料集共同欄位，紀錄每個資料的 ID、性別和年齡等基本資料，以及一些生理指標、健康狀況和心理健康與症狀。

表一、資料集共同欄位

人口統計學	生理指標	健康狀況	心理健康與症狀
Sub_ID 唯一識別碼 sex age Ethnicity 種族 Marital_status 婚姻 Handedness_category 慣用手 Education_years	height weight systolic 收縮壓 Diastolic 舒張壓 pulse 脈搏	hypertension 高血壓 diabetes 糖尿病 stroke 中風 CVD 心血管疾病 TBI 創傷性腦損傷 cancer 癌症 Neurological_diseases 神經系統疾病 Other_diseases 其他疾病	多種心理評估量表 (HAM-D, HAM-A, PANSS 等) 睡眠質量評估 (PSQI 系列) 家族精神病史 (I/II/III-family_psychiatric_history)

在三個精神疾病中，思覺失調症、重鬱症和躁鬱症有一些部分共同欄位，如表二，另外思覺失調症和躁鬱症各擁有一個專有欄位。思覺失調症中，Family_code 表示受試者家屬是否也包含在 TAMI 中的 ID 編號(MRSF001 至 MRSF240)；躁鬱症的Type_of_diagnosis 則表示雙相情感障礙 I 型或 II 型。

表二、思覺失調症、重鬱症和躁鬱症部分共同欄位

躁鬱症	重鬱症	思覺失調症
Onset_age (發病年齡)		
Medication (藥物類型) I-family_psychiatric_history II-family_psychiatric_history III-family_psychiatric_history		
	Date_of_birth Date_of_scan (MRI) I-family_psychiatric_type II-family_psychiatric_type III-family_psychiatric_type	

所有受試者資料，不同 ID 皆放有相同資料夾，內容如下：

1. Touch：用於標記特定處理步驟的完成狀態，記錄多種大腦影像處理任務的進展。
2. Surf：包含與大腦皮層表面模型和相關統計數據有關的檔案。
3. Scripts：與 FreeSurfer 腦部影像重建流程相關的文件。
4. MRI：包含了處理後的各種 MRI 結果影像。
5. Stats：支援量化分析，允許比較不同受試者或群體的腦部形態特徵。
6. Label：腦區分割的定義、視覺化和分析提供基礎。

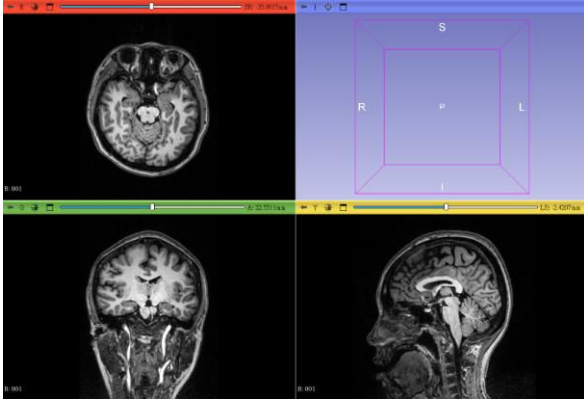
在健康組和三個精神疾病（思覺失調症、重鬱症和躁鬱症）的資料集中，各有一個 Excel 存放每個受試者各個腦區的異常情況，並做內、外部質量檢查。分析後發現年齡越高越容易出現所有腦區都左、右半腦都異常的情況，並且找出表四的各精神疾病與大腦區域關聯。

表三、各精神疾病與大腦區域關聯

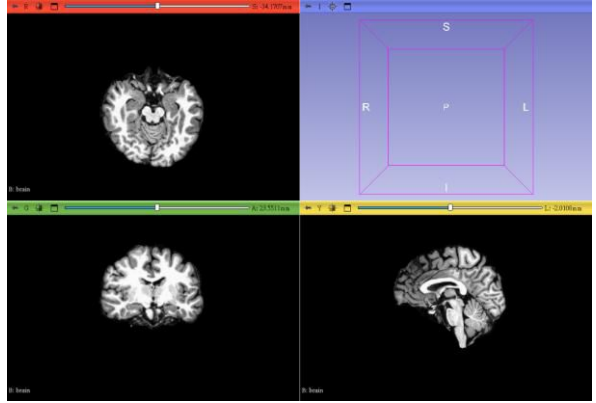
大腦區域名稱	思覺失調症	重鬱症	躁鬱症
Superior temporal 顳上溝			顯著表現
Supramarginal 上緣回			顯著表現
Insula 島葉	較強表現		顯著表現
Temporal pole 顳葉端			較強表現
Precentral 中央前回			較強表現
Caudalmiddlefrontal 尾側中額葉區			較強表現
Pars opercularis 上顳迴區與額下迴上區		較強表現	較強表現
Cuneus 楔葉	顯著表現	顯著表現	
Lingual 舌回		顯著表現	
Pericalcarine 距狀裂周圍皮質		較強表現	
Superiorfrontal 上額葉皮質	較強表現		
Superior parietal 上頂葉皮質	顯著表現		
Medial orbitofrontal 內側眶額皮質	顯著表現		

4.2 001.mgz 與 brain.mgz

本研究中使用兩種影像，分別是涵蓋完整頭部的影像 001.mgz，以及去除顱骨後的大腦影像 brain.mgz。使用 3D Slicer 開源軟體匯入 .mgz 檔案，顯示 MRI 腦部 3D 立體影像圖，如下。



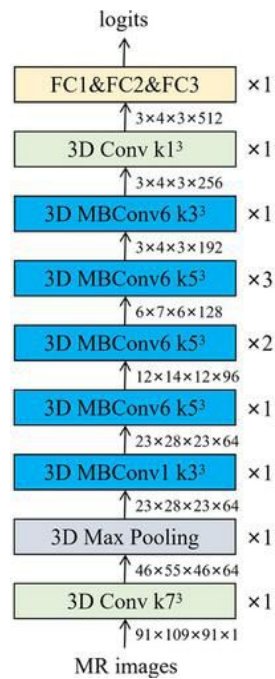
圖五、001.mgz



圖六、brain.mgz

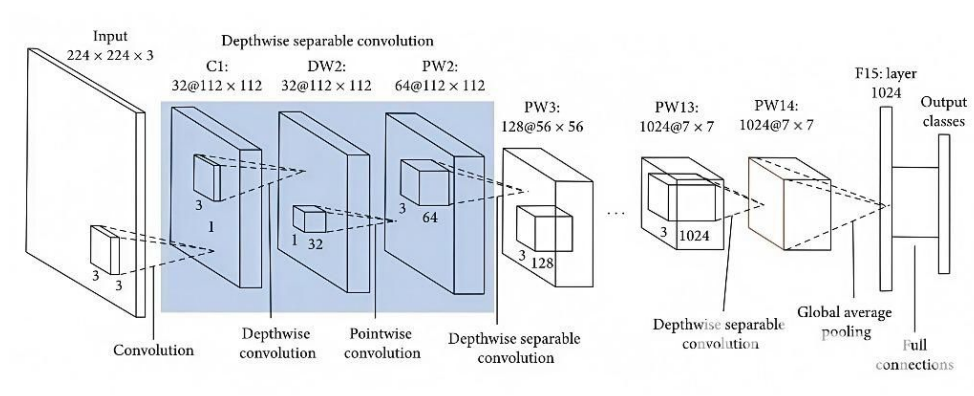
4.3 3D EfficientNet 模型架構

3D EfficientNet 使用複合縮放方法來擴展網路維度，以固定比例來均勻縮放寬度、深度和解析度。可用於處理有體積和時間維度的資料，具有較高的效能和準確率。圖四為 3D EfficientNet 架構圖[11]。3D convolution layer（3D 卷積層）和 3D pooling layer（3D 池化層）降低特徵圖的空間尺度，並提取 MRI 影像中的特徵，接著由五個 3D MBConv 區塊處理特徵圖[11]。



圖七、3D EfficientNet 架構圖[11]

其相關核心架構採用深度可分離卷積 (Depthwise Separable Convolution)，包括深度卷積 (Depthwise) 和逐點卷積 (Pointwise)，大幅減少參數量和計算成本，同時維持高準確率。最後透過全局平均池化 (Global Average Pooling) 和全連接層輸出分類結果。



圖八、EfficientNet 架構圖[12]

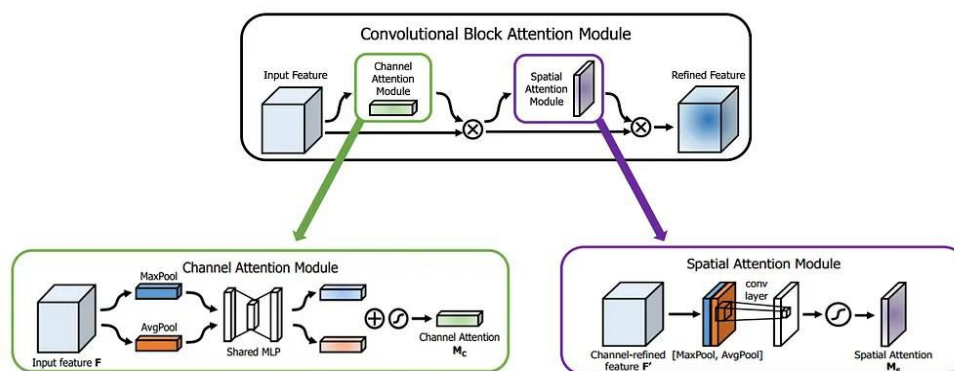
本研究採用 Shijian Jian 等人開源的 3D EfficientNet 架構為基礎 [13]，設計並訓練多項精神疾病影像分類模型，涵蓋健康受試者與三種精神疾病（躁鬱症、重鬱症和思覺失調症）之間的不同分組。為確保模型評估的穩定性與泛化能力，本研究使用 20-Fold Cross Validation（20 折交叉驗證）進行訓練與驗證。

模型訓練與測試流程使用各類別 100 筆影像資料作為訓練基礎，在每個 epoch 結束後，會輸出訓練與驗證準確率，並儲存模型權重，以利於在訓練中斷後能接續進行，確保實驗結果可重現。

此外，本研究針對「健康與精神疾病 CBAM 二元分類」，進一步在 3D EfficientNet 架構中整合 Convolutional Block Attention Module (CBAM)。此模組透過結合通道注意力 (Channel Attention) 與空間注意力 (Spatial Attention) 機制，強化模型對腦部影像中關鍵區域的特徵捕捉能力，進而提升分類準確率與模型對異質性資料的辨識表現。

4.4 CBAM

通道空間注意力機制 (Convolutional Block Attention Module, CBAM) 用於提升模型對於影像特徵的識別能力，是由通道注意力模組 (Channel Attention Module, CAM) 和空間注意力模組 (Spatial Attention Module, SAM) 所組成。道通注意力模組透過 Max Pooling 和 Average Pooling 得到兩個不同空間的表徵向量，再使用 MLP 進行降維，最後將兩個特徵向量相加並經過 Sigmoid 得到最後的通道域注意力向量[14]。空間注意力模組則是透過 Max Pooling 和 Average Pooling 得到兩個不同的空間的特徵圖，將兩個疊加在一起並經過 Conv Layer 和 sigmoid 激發函數，最後得到空間注意力特徵圖[14]。



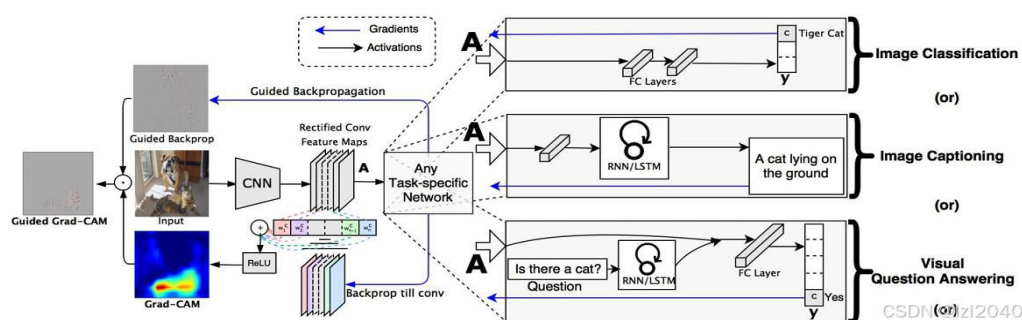
圖九、CBAM 架構圖[14]

4.5 Grad-CAM

Grad-CAM (Gradient-weighted Class Activation Mapping) 是一種深度學習模型的可解釋性方法，主要透過分析輸出類別相對於卷積層特徵圖的梯度，計算各特徵圖的重要程度，進而產生類別權重的活化熱力圖。此熱力圖顯示模型在進行分類判斷時所關注的主要影像區域，使研究者了解模型決策是否基於具有意義的局部特徵[15]。

本研究將 Grad-CAM 結合 3D EfficientNet 模型，在健康與疾病之二分類模型完成訓練後，對不同輸入影像生成熱力圖，並將其疊加於原始 MRI 影像之上，展示模型在

做出預測時所聚焦的位置。藉由觀察 Grad-CAM 結果的視覺化分析，不僅能確認模型是否成功捕捉與病灶或健康相關的重要區域，也能進一步評估模型的可靠性與可解釋性，並作為後續模型調整與改善的參考。



圖十、 Grad-CAM 架構圖[16]

4.6 K-Fold Cross Validation (交叉驗證)

為了提升模型評估的穩定性與泛化能力，本研究採用 K-Fold Cross Validation 方法進行訓練與驗證。此方法能在有限的資料量下，最大化資料利用率，並避免因單一資料切分造成的模型偏差。本研究設定 $K=20$ ，即將整體資料分為 20 組，於每次訓練中輪流選取 1 份作為測試集，其餘 19 份作為訓練與驗證資料，最終取各結果的平均值作為整體模型的最終表現。



圖十一、K Fold 工作流程[17]

4.7 Learning Rate 設定與學習率調度器 (Scheduler)

1. 初始設定：

- 優化器：Adam
- 初始學習率 (LR)：0.0001
- 權重衰減 (weight decay)：1e-5

2. 學習率調整方式：

採用 StepLR 調度器，每訓練 10 個 epoch 將學習率乘上 0.5。

3. 設定目的：

- 初期較高的學習率可加快收斂。
- 中後期降低學習率能穩定模型權重更新。

Chapter 5 研究成果

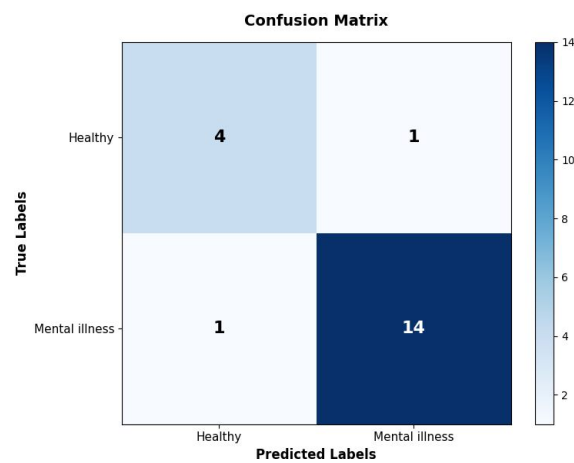
5.1 健康與精神疾病 CBAM 二元分類

目前使用 K-Fold($K = 20$) 並加入 CBAM 通道空間注意力機制來訓練健康與精神疾病 CBAM 二元分類，健康的資料集筆數為 100 筆，而精神疾病包含躁鬱症、重鬱症和思覺失調症各 100 筆，其中 Fold 19 份訓練與驗證資料以 80% 訓練集比 20% 驗證集分割，已經達到 Average Test Accuracy 74.25%，並且表五中 Precision、Recall 和 F1 Score 也皆達 82% 以上。

表四、各精神疾病與大腦區域關聯

指標	數據
Average Test Accuracy	74.25% $\pm$ 0.1228
Precision	82.66% $\pm$ 0.0705
Recall	82.67% $\pm$ 0.1511
F1 Score	82.03% $\pm$ 0.1160

在健康與精神疾病二元分類 K-Fold(K = 20) 中，在 Fold 15 有 Best Test Accuracy 90.00%，健康和精神疾病各錯一筆資料，如圖七的混淆矩陣，健康的指標皆為 80%，而精神疾病的指標皆為 93.33%，如表六。因此，健康與精神疾病 CBAM 二元分類可以透過 MRI 影像有效分辨患者是否患有精神疾病。



圖十二、Fold 15 混淆矩陣

表五、Fold 15 指標

	Precision	Recall	F1 Score
健康	80.00%	80.00%	80.00%
精神疾病	93.33%	93.33%	93.33%

5.2 年齡預測

目前已成功使用健康組的 24 筆 orig.mgz 的 MRI 影像數據訓練模型，其中包括年齡為 24、29、36、59、60、62、64 和 74 的各三筆資料，在 Google Colab 雲端開發環境成功訓練 3DEfficientNet 模型，如圖五。隨著模型訓練次數增加，Loss 穩定下降，Accuracy 也持續上升，最終模型準確率達到 87.50%，代表模型已趨於收斂並具穩定性。預測驗證資料的平均誤差為 7 歲，一筆實際年齡 29 歲的資料，會被預測為 36 歲。未來將繼續新增資料集數量進行訓練，以增加預測的精準度。

```
Epoch [1] Loss: 4.558 Accuracy: 8.33%
Epoch [2] Loss: 4.159 Accuracy: 20.83%
Epoch [3] Loss: 3.813 Accuracy: 29.17%
Epoch [4] Loss: 3.462 Accuracy: 41.67%
Epoch [5] Loss: 3.212 Accuracy: 29.17%
Epoch [6] Loss: 2.998 Accuracy: 45.83%
Epoch [7] Loss: 2.779 Accuracy: 37.50%
Epoch [8] Loss: 2.582 Accuracy: 37.50%
Epoch [9] Loss: 2.351 Accuracy: 58.33%
Epoch [10] Loss: 2.358 Accuracy: 54.17%
Epoch [11] Loss: 2.021 Accuracy: 66.67%
Epoch [12] Loss: 1.950 Accuracy: 66.67%
Epoch [13] Loss: 1.971 Accuracy: 70.83%
Epoch [14] Loss: 1.834 Accuracy: 87.50%
Epoch [15] Loss: 1.760 Accuracy: 83.33%
Epoch [16] Loss: 1.725 Accuracy: 70.83%
Epoch [17] Loss: 1.620 Accuracy: 70.83%
Epoch [18] Loss: 1.506 Accuracy: 87.50%
Epoch [19] Loss: 1.560 Accuracy: 87.50%
Epoch [20] Loss: 1.692 Accuracy: 87.50%
Finished Training
```

圖十三、3D EfficientNet 模型訓練

5.3 實驗流程

H: Healthy

B: Bipolar Disorder

M: MDD

S: Schizophrenia

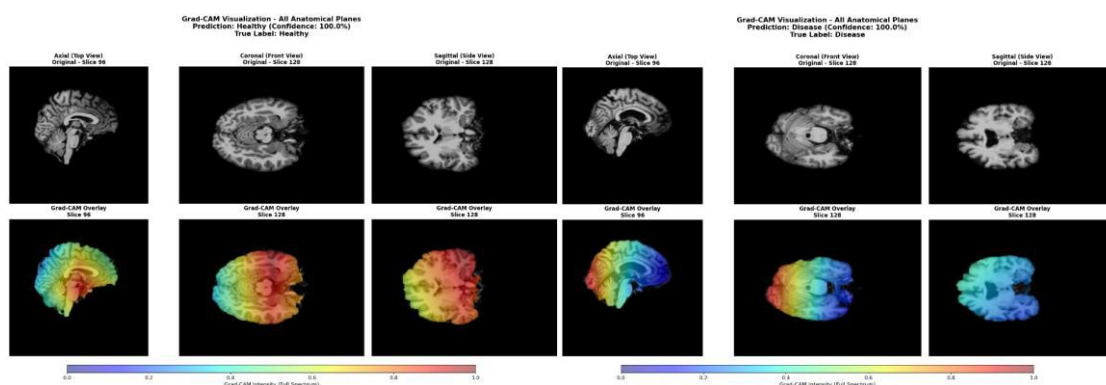
表六、實驗流程

類型	實驗名稱	分類組合	測試準確率
四分類	健康與三種精神疾病分類	H vs B vs M vs S	48%
三分類	三種精神疾病分類	B vs M vs S	38%
二分類	健康與精神疾病 二元分類	H vs (B, M, S)	70%
	一種精神疾病對 健康及兩種精神疾病分類	B vs (H, M, S) M vs (H, B, S) S vs (H, B, M)	71% 71% 63%
	一種精神疾病及健康對 兩種精神疾病分類	(B, H) vs (M, S) (M, H) vs (B, S)	53% 45%
	一種精神疾病對健康	B vs H M vs H S vs H	70% 60% 70%
	一種精神疾病對 兩種精神疾病	B vs (M, S) M vs (B, S) S vs (B, M)	62% 62% 63%
	一種精神疾病對 一種精神疾病	B vs M B vs S M vs S	55% 50% 50%
四分類	健康與精神疾病二元分類 + CBAM	H vs (B, M, S)	74.25%

Chapter 6 實驗結果分析

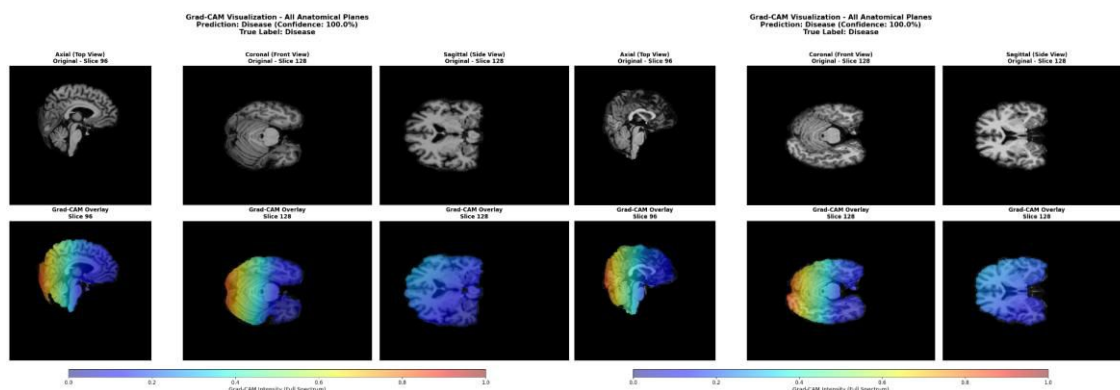
本研究採用 3D EfficientNet 完成健康與精神疾病 CBAM 二元分類訓練，其中於第 15 折 (Fold 15) 獲得最高測試準確率 Test Acc: 90.00%。為進一步檢視模型的判斷依據，我們對該折表現最佳之模型進行 Grad-CAM 視覺化，並針對其分別在健康與三種疾病案例的不同解剖平面 (Axial、Coronal、Sagittal) 關注區域進行分析。結果顯示，模型在兩類別上分別呈現一致特徵，使其具有可解釋性。

在健康樣本中，熱力圖多集中於大腦皮質與中線結構附近，尤其可見沿視丘及兩側半球皮質呈現較高關注。此分布可能反映模型將皮質對稱性與白質邊界清晰度作為健康判讀依據。相較之下，疾病樣本的 Grad-CAM 關注區域多集中於後側腦部 (posterior region)。此結果顯示模型在進行疾病分類時，可能偵測到後側皮質與鄰近部位之結構差異，並將其視為疾病判定的主要依據。



圖十四、Grad-CAM 健康樣本

圖十五、Grad-CAM 思覺失調症樣本



圖十六、Grad-CAM 躁鬱症樣本

圖十七、Grad-CAM 重度憂鬱症樣本

Chapter 7 結論

在這項研究中，我們成功以客觀的影像學分析來進行精神疾病診斷，目前已經能透過腦部 MRI 影像，有效分辨患者是否患有精神疾病，健康與精神疾病 CBAM 二元分類使用 K-Fold($K = 20$) 的 Average Test Accuracy 已經達到 74.25%，且從 Grad-CAM 可看出健康與三種精神疾病有不同的關注區域。然而經過多種實驗組合，探討三種精神疾病的關聯性，發現以影像來分析仍有難度，無法有效區分三者之間的差異。

我們期望提升精神疾病的早期篩檢效率，並持續提升模型的判斷準確率，建立可實際應用於臨床的精神疾病影像辨識輔助診斷系統。

文獻參考

- [1] 揚生慈善基金會 Understanding mental health: Key insights.
<https://www.ysfoundation.org.tw/post.aspx?Id=1082&Page=1>
- [2] World Health Organization. (2022). *Mental disorders*. <https://www.who.int/news-room/fact-sheets/detail/mental-disorders>
- [3] Christina A. Samuels. (2013, June 3). Disability definitions revised in psychiatric manual. Education Week. <https://www.edweek.org/leadership/disability-definitions-revised-in-psychiatric-manual/2013/06?tkn=LNUFsPdoPW0PIHMGdelm%2Bi8FZjx3sl4FAJLJ&cmp=clp-edweek>
- [4] National Science and Technology Museum. (n.d.). Exploring mental health and technology. <https://scitechvista.nat.gov.tw/Article/C000003/detail?ID=6ce56dc1-e7ac-4ea9-b2e1-d00a41bd46f3>
- [5] 報導者 Taiwan Reporter. (2025). Taiwan citizen judge case: A boy's death in nanny care. <https://www.twreporter.org/a/taiwan-citizen-judge-a-boy-death-in-nanny-care-2>
- [6] Reangel. (n.d.). Mental health awareness and support.
<https://www.reangel.com/Article.php?LI=320>
- [7] 東台灣新聞網 eTaiwan News. (2022, June 27). Depression should not be ignored: Complete disclosure of symptoms to avoid misdiagnosis.
<https://etaiwan.news/2022/06/27/躁鬱症不可輕忽-完整交待症狀避免醫師誤判為/>
- [8] Albert C. Yang, Chu-Chung Huang, Heng-Liang Yeh, Mu-En Liu, Chen-Jee Hong, Pei-Chi Tu, Jin-Fan Chen, Norden E. Huang, Chung-Kang Peng, Ching-Po Lin, Shih-Jen Tsai. Complexity of spontaneous BOLD activity in default mode network is correlated with

- cognitive function in normal male elderly: a multiscale entropy analysis. Volume 34, Issue 2, Pages 428-438. (2013)
- [9] Zhu JD, Wu YF, Tsai SJ, Lin CP, Yang AC. Investigating brain aging trajectory deviations in different brain regions of individuals with schizophrenia using multimodal magnetic resonance imaging and brain-age prediction: a multicenter study. doi: 10.1038/s41398-023-02379-5. (2023)
- [10] National Yang Ming Chiao Tung University. (n.d.). Taiwan Aging and Mental Illness Cohort (TAMI). DMC Data Repository. <https://data.dmc.nycu.edu.tw/dataset/taiwan-aging-and-mental-illness-cohort-tami>
- [11] Bowen Zheng, Ang Gao, Xiaona Huang, Yuhan Li, Dong Liang, Xiaojing Long. A modified 3D EfficientNet for the classification of Alzheimer's disease using structural magnetic resonance images. IET Image Processing, (2022). 16(10), 2427–2438. <https://doi.org/10.1049/ipr2.12618>.
- [12] Mingxing Tan, Quoc V. Le, EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks [ICML 2019]
- [13] S. Jian. EfficientNet-PyTorch-3D. GitHub. <https://github.com/shijianjian/EfficientNet-PyTorch-3D>
- [14] Sanghyun Woo, Jongchan Park, Joon-Young Lee, In So Kweon. CBAM: Convolutional Block Attention Module. arXiv (2018). <https://doi.org/10.48550/arXiv.1807.06521>.
- [15] iT 邦幫忙技術文章. *Grad-CAM*：深度學習模型的視覺解釋方法. iT 邦幫忙. <https://ithelp.ithome.com.tw/m/articles/10334625>
- [16] CSDN 技術部落格. *Grad-CAM* 熱力圖實作與示例. CSDN. https://blog.csdn.net/qq_41234663/article/details/139537020
- [17] cross-validation：從 holdout validation 到 k-fold validation. <https://blog.csdn.net/lanchunhui/article/details/50522424>

附錄 A. 專題工作內容

鄭又瑄、劉喬恩、郭喬恩

我們於郭文嘉老師指導下進行專題研究，題目為「應用深度學習於精神疾病腦部 MRI 影像分析」。本專題結合神經影像處理與深度學習技術，利用台灣健康老化與精神疾病腦影像資料庫 (TAMI) 進行 MRI 腦部影像分析，目標為精神疾病分類模型，協助開發具臨床潛力之診斷輔助系統。

本研究同時處理結構影像分割與分類任務，主要工作可分為以下幾項：

- 資料理解與資料前準備
 - 採用 TAMI 資料庫中健康組與三類精神疾病（思覺失調症、重鬱症、躁鬱症）共 1,200 多筆 MRI 影像，並整理各組 Excel 資料中的腦區異常資訊作為輔助分析依據。
 - 使用涵蓋完整頭部的影像 001.mgz，以及去除顱骨後的大腦影像 brain.mgz。
 - 皆維持原始 MRI 影像體素矩陣大小 $192 \times 256 \times 256$ 的解析度。
 - 匯整並統計各疾病與特定腦區間的對應關係。
- 模型訓練
 - 使用 Conda 安裝 python 虛擬環境，並透過 3D EfficientNet 模型進行精神疾病腦部特徵分析。
 - 儲存各階段模型與權重，便於後續持續訓練與分析。
- 精神疾病分析與視覺化
 - 探討健康與三種精神疾病的多種組合。
 - 探討三種精神疾病間的多種組合。

- 導入 Grad-CAM 熱區圖進行模型可視化分析，以觀察模型在分類過程中關注的腦區，並對照實際異常結構。
- 成果彙整與未來方向
 - 撰寫期中與期末報告，完整記錄專題進行過程與分析結果。
 - 依據指導老師建議，規劃後續擴展內容。

附錄 B. 專題心得與建議

◆ 鄭又瑄

猶記最初剛訂定專題主題時，我們曾因對醫療知識陌生擔憂後續研究。但隨著深入研究和報告 TAMI 資料集，逐步吸收了腦部基本構造和精神疾病相關區域的知識。過程中我特別有感，應用醫療知識於訓練模型，即使並非專業，只要善用網路資源和科技工具，也能逐步克服知識門檻，慢慢理解學習進而消弭擔憂。

而進行專題研究，最初是使用 Google Colab 試圖透過程式碼讀取與分析 .mgz 檔案，程式碼跟用到 python 以前從未接觸過的函式庫，後來還發現開源套件 3D Slicer 可直接顯示三軸切面及 3D 立體成像，操作時感到新奇和實用。雖與程式無關，但勾畫腦部 MRI 的過程中，視覺化功能讓我清楚辨識腦區細節。

暑假過後，我們開始使用虛擬環境執行模型訓練，我認為團隊運作之所以順利，關鍵在於每週會議與組內討論確保分工明確。雖然研究的三種疾病之間並不容易區分特徵，但疾病與健康類別樣本的差異可以從熱力圖清楚的視覺化結果。在專題實作期間，我不僅學會建置 python 環境，也完整了解醫療影像的機器學習過程，更熟悉熱力圖的實作方式。

未來建議：

- 做專題時每週檢視進展，並提前規劃進度，後續再動態調整，較可以確保階段性的任務按時完成。
- 於研究初期評估資料品質與可區分性，優先選擇標註清楚、特徵差異明確的問題與資料集，提升模型訓練的穩定性與研究可行性。

◆ 劉喬恩

本次專題研究，我學到很多關於 3D EfficientNet 模型的建置、訓練和測試，以及將 K Fold 和 CBAM 加入模型的用法，還有分析資料集的方法。

從一開始完全不會任何模型，所以我會先去網路上找很多資料，了解模型的架構和功能，也會觀看一些教學影片，學習模型的建構方式。在建構模型的過程會碰到很多困難，像是 Colab 版本和 mgz 影像大小等許多問題，在解決這些問題的過程又會學到不少模型建構需要注意的地方。另外我也學會可以撰寫程式儲存每個 epoch 訓練完後的模型和權重，才可以在執行中斷後不用重新訓練，否則可能需要再重新花一、兩個小時去跑模型。還有了解 K Fold 和 CBAM 的架構，並學會如何使用在模型當中，提升模型的準確率。

其實分析資料集沒有想像中容易，因為裡面會有好幾百筆資料，每個資料內又有很多不同的檔案，需要很仔細的去找出檔案間的關聯，而且因為我本身不是醫學背景，不了解任何醫學專業知識，檔案裡面的東西，需要從頭開始學起，一丁一點的自行學習關於精神疾病和 mgz 腦部影像等的醫學知識。最後在健康與精神疾病 CBAM 二元分類取得 Average Test Accuracy 74.25%，能夠為精神疾病領域做出貢獻，期許未來這方面的研究能夠更進步，實際運用在醫學上幫助他人。

未來建議：

- 專題前或初期可以學習模型相關課程
- 專題初期瞭解精神疾病的相關醫學知識
- 隨時關注精神疾病辨識的相關研究

◆ 郭喬恩

這次的專題是我第一次嘗試將程式、醫學影像與機器學習結合在一起，對我來說是一個全新的挑戰。一開始接觸 TAMI 資料集時，面對眾多不熟悉的檔案格式（如 .mgz）、命名方式與資料架構，確實讓我感到有些混亂。但隨著逐步摸索，透過查看說明文件、嘗試寫簡單的讀檔程式、參考開源工具的操作方式，逐漸建立起對醫學影像資料的基本理解，也提升了我的資料處理能力。

在專題進行的過程中，也讓我認識到與組員之間協作的重要性。我們在分工上逐漸找到適合各自的角色，也會利用雲端同步進度。雖然有時候遇到卡關，也曾經出現方向不夠清晰的狀況，但透過討論與資料查找，問題總能找到出路。這種「一起解決困難」的經驗，讓我對團隊合作有更深入的理解。

回顧目前的進展，我認為未來可以進一步加強跨領域知識的整合能力。醫學影像分析不僅需要程式技巧，還需要基本的神經科學背景、影像處理常識，以及對病理的初步理解。這些對我們非醫學背景的學生而言是很大的挑戰，但若能提早安排學習順序與重點，例如先理解腦部主要結構與常見異常類型，再進行建模，會更有效率。

未來建議：

- 導入版本控制工具（如 Git）：方便多人協作與程式碼管理，也能更清楚每次修改的內容與進度。
- 事前與期中定期設立 milestone：避免中後期才發現進度落後，應依任務階段明確設立小目標，有利於推進整體進度。
- 多參考前人實作案例：如 kaggle、GitHub 的開源專案，不僅能學習技術細節，也能參考其專題架構與報告方式。