

NKUST 機器學習期末報告

隊伍：TEAM_9317

隊員：羅祥恩

Private leaderboard：0.931769 / Rank 229



壹、資料處理與程式環境

資料前處理

本研究參與 AI CUP 2025 競賽二「主動脈瓣膜物件偵測」，任務目標為利用深度學習模型自醫學影像中準確偵測主動脈瓣膜（Aortic Valve）位置。本研究使用之資料集由競賽官方提供，包含 training_image 與 training_label 兩部分，並透過 Google Drive 下載後進行解壓與整理。影像資料主要為醫學影像格式，標註資料則已轉換為 YOLO 所需之標註格式。

在資料前處理階段，首先確認影像與標註檔案之一致性，避免出現影像存在但無對應標註，或標註檔案錯誤之情況。接著，檢查標註框是否超出影像邊界，並修正或移除異常標註資料，以確保訓練資料品質。為提升模型泛化能力，本研究亦在訓練過程中啟用 YOLO 內建之資料增強機制，例如隨機翻轉、縮放、亮度調整與 Mosaic augmentation，使模型能學習不同尺度與角度下之主動脈瓣膜特徵。

最後，依據競賽設定與訓練需求，透過 `aortic_valve_colab.yaml` 設定檔指定訓練集與驗證集路徑、類別名稱與類別數量，確保模型訓練流程可順利進行。

原始程式碼概述

本研究之程式碼主要以 Python 撰寫，並採用模組化設計，整體架構可分為資料下載與前處理、模型訓練、模型驗證與預測三大部分。資料下載部分利用 `gdown` 套件自 Google Drive 下載官方資料集與設定檔；模型訓練部分則呼叫 YOLOv12 框架進行模型初始化與訓練；模型驗證與預測部分負責將訓練完成之模型套用於驗證資料集，並輸出預測結果供後續分析。

透過清楚的程式架構與註解說明，使程式具備良好可讀性，並方便第三方使用者依照說明進行環境建置、重新訓練與結果重現。

執行環境

本研究之實驗環境為雲端 Colab 平台（Linux 作業系統），程式語言採用 Python 3。主要使用之函式庫包含 PyTorch、Ultralytics YOLOv12、NumPy、OpenCV 與 Matplotlib。模型訓練過程中使用 GPU（NVIDIA Tesla 系列）進行加速，以縮短訓練時間並提升訓練效率。

本研究使用 YOLOv12 官方提供之預訓練權重作為初始模型，透過遷移學習方式使模型能快速學習主動脈瓣膜之特徵。預訓練模型在初期已具備基本影像特徵擷取能力，有效提升模型收斂速度與整體偵測效能。

貳、模型設計與訓練策略

模型訓練與預測原始程式碼與說明

本研究採用 YOLOv12 作為主要物件偵測模型，並將完整訓練、驗證與預測程式碼上傳至 GitHub 平台，供第三方使用者參考與重現實驗結果。GitHub 專案中包含資料下載指令、模型訓練流程、驗證與推論程式碼，並附有 README.md 檔案詳細說明環境建置方式、套件需求與執行步驟。

程式碼內容涵蓋資料處理、模型訓練、權重儲存與模型推論等功能，使用者可依說明快速完成模型重訓或套用模型進行預測。

Github 連結：[GitHub - c112154112-sys/Machine-Learning](https://github.com/c112154112-sys/Machine-Learning): 機器學習期末報告

模型訓練流程

模型訓練流程首先載入 `aortic_valve_colab.yaml` 設定檔，確認資料路徑與類別設定正確後，初始化 YOLOv12 模型與預訓練權重。訓練過程採用監督式學習方式，透過反向傳播演算法不斷更新模型權重，以最小化偵測損失函數。

在每個訓練 epoch 結束後，模型會使用驗證資料集進行效能評估，以監控模型是否出現過度擬合現象。當驗證結果達到最佳表現時，即保存該模型權重作為最終模型，用於後續測試與分析。

參數設定

本研究模型訓練之主要參數包含訓練次數（epochs）、批次大小（batch size）、學習率（learning rate）與影像輸入尺寸（image size）。訓練次數依據模型收斂情況設定，批次大小則依 GPU 記憶體限制進行調整。學習率採用 YOLO 預設設定，以兼顧訓練穩定性與收斂速度。上述參數設定經多次實驗調整後，能有效提升主動脈瓣膜之偵測準確率。

參、分析與結論

本研究使用 YOLOv12 模型應用於 AI CUP 2025 主動脈瓣膜物件偵測競賽，實驗結果顯示模型能有效定位主動脈瓣膜位置，並在驗證資料集上取得穩定的偵測表現。透過資料前處理與資料增強策略，使模型能適應不同影像條件與變化，提升整體泛化能力。

然而，模型在部分影像中仍可能出現誤判或漏判情形，可能原因包含影像品質不佳、瓣膜邊界模糊或資料數量有限等因素。未來可考慮引入更多訓練資料、調整模型深度或嘗試不同的資料增強策略，以進一步改善模型效能。此外，也可比較不同 YOLO 版本或其他物件偵測模型之表現，以尋找更適合醫學影像偵測任務之方法。

整體而言，YOLOv12 搭配預訓練權重與適當的訓練策略，在主動脈瓣膜物件偵測任務中展現良好潛力，並證實深度學習技術於醫學影像分析領域之應用價值。

此外，從競賽實務角度來看，主動脈瓣膜屬於醫學影像中結構相對細小且邊界不明顯之目標，對物件偵測模型而言具有一定挑戰性。本研究透過 YOLOv12 即時物件偵測架構，成功在準確率與推論速度之間取得平衡，顯示其於醫學影像輔助診斷應用上的可行性。未來若能結合更多醫學專業標註資料，或導入多尺度特徵融合與注意力機制，將有助於進一步提升模型對細微結構的辨識能力，並增強其在臨床應用場景中的實用價值。

肆、使用的外部資源與參考文獻

Ultralytics. (2024). YOLOv12 documentation. Ultralytics. <https://github.com/ultralytics>

AI CUP. (2025). AI CUP 2025 競賽二：主動脈瓣膜物件偵測. <https://aicup.tw>

OpenAI. (2024). ChatGPT (GPT-4) [Large language model]. <https://chat.openai.com>