書報討論書面報告

11363117 戴子璿

日期:2025/04/01

講者:廖弘源 博士

題目:Yolov4,v7,v9的發展過程

心得報告:

Yolo已經是很多應用開發會使用到的影像辨識套件，今天也聽到了廖博士在開發這AI的心路歷程，也聽到了很多調整AI的方法。

講到影像辨識，就不能不提到視覺字典，ImageNet 是一個非常大型的圖像資料庫，ImageNet 就像是一個「詞語到圖像」的視覺字典，使電腦可以透過視覺字典查詢狗是甚麼，飛機是甚麼。

在演講中也提到封閉曲線，這解釋了電腦如何辨識圖形、形狀與物體，尤其是在邊界資訊與輪廓結構方面，在圖像中，封閉曲線（Closed Curve）指的是一條首尾相連的連續曲線，有助於區分圖形與背景，幫助人眼快速辨識物體，也可以幫助電腦可以輪廓提取、形狀分類、物體分割、邊界學習等。

在這堂課中學習到了一個很重要的電腦視覺目標偵測方法，R-NN two-stage object detectors。Two-stage目標偵測器先找出可能位置，再分類並微調的模型。Region Proposal先找出圖片中可能有物體的位置，Object Classification & Bounding Box Refinement，再對每個 proposal 做分類（是什麼）和框的位置微調（更精確定位），而這樣的方式，與one-stage的方式準確率更高，但相對執行速度較慢。

這次的課程介紹了很多影像辨識領域的概念，不只知道YOLO廣泛的應用方式，也了解了廖博士在開發時的經驗分享及挑整方法。

參考文獻: K. Jannet, M. S. Sunar, M. M. Islam Molla and M. A. Bin As'Ari, "A Deep Learning Approach to Badminton Player Footwork Detection Based on YOLO Models: A Comparative Study," 2024 IEEE 8th International Conference on Signal and Image Processing Applications (ICSIPA), Kuala Lumpur, Malaysia, 2024, pp. 1-6, doi: 10.1109/ICSIPA62061.2024.10686537.

這篇論文透過YOLO模型來偵測羽球選手的步伐動作，來協助訓練與比賽分系，比較了YOLOv8與YOLOv9，發現YOLOv8的表現最好，而最終也可以準確地偵測選手動作位置，實現實時分析與教練訓練回饋。