

# AIMBOT

專題編號：114-2-CSIE-S007

執行期限：114 年第 1 學期至 114 年第 2 學期

指導教授：陳彥霖

專題參與人員： 112590021 張家程  
112590023 林宇祥  
112590055 彭瑜祐

## 一、摘要

本作品為 AIMBOT 即時人物偵測與雲台追蹤系統，旨在整合人物偵測、多目標追蹤、重識別與伺服控制，建立可即時鎖定並持續跟隨指定目標之智慧化影像伺服平台。系統以攝影機影像串流為輸入，採用 YOLOv7 進行人物偵測，透過 BoT-SORT 作為主要多目標追蹤框架，必要時並以 ByteTrack 作為原始對照，以維持多目標 ID 穩定追蹤。系統並使用 OSNet x0\_5 強化重識別能力；在目標短暫遮擋或離開視野後，利用 OSNet 特徵比對與 lost-track recovery 機制進行重識別回復。控制端則以 PID 演算法將畫面誤差轉換為雲台角速度，以達成目標居中與持續追蹤。為提升系統即時性，另整合影格預取、GPU 預處理與非同步推論架構。實作結果顯示，本作品已完成核心模組整合，具備即時追蹤、目標回復與互動式操作介面等功能，可應用於智慧監控、自動導引與視覺追蹤等領域。

**關鍵詞：**YOLOv7、BoT-SORT、OSNet x0\_5、PID 控制、即時影像追蹤、雲台控制

## 二、緣由與目的

近年來，電腦視覺、深度學習與智慧控制技術快速發展，使即時目標偵測與追蹤系統在智慧監控、自主設備、人機互動及視覺導引等領域具有高度應用價值。然

而，實際系統除了需具備良好的偵測準確率外，尚須兼顧多目標追蹤穩定性、目標短暫遮擋後之回復能力，以及控制端對目標偏移的即時修正能力。

本作品 AIMBOT 之設計目標，在於整合 YOLOv7、BoT-SORT、OSNet x0\_5 與 PID 控制方法，建立一套可由使用者指定追蹤目標並驅動雲台自動跟隨之即時系統，藉以驗證模組化視覺追蹤架構在即時性、穩定性與擴充性上的可行性，並以 ByteTrack 作為原始對照，評估不同追蹤策略於目標關聯與重識別整合上的差異，作為智慧影像伺服系統之實作成果展示。

## 三、研究報告內容

本作品之系統架構包含影像輸入、人物偵測、多目標追蹤、重識別、控制演算法及互動介面等模組。其流程為：首先由攝影機擷取影像，透過 FramePrefetcher 與 GPU 預處理模組完成影格快取與縮放；其次以 YOLOv7 進行人物偵測，再由 BoT-SORT 作為主要追蹤器以維持目標 ID 穩定，必要時以 ByteTrack 作為原始對照比較追蹤表現；當目標因遮擋或離開視野而暫時失聯時，啟用 OSNet x0\_5 進行特徵擷取，並透過 OSNet 特徵比對與 lost-track recovery 機制回復原有目標識別。最後，控制模組依據目標中心與畫面中心之誤差，透過每軸 PID 將偏差轉換為雲台角速度，完成目標居中與持續跟隨。

#### 四、系統實作成果與分析

在使用技術與工具方面，本作品以 Python 為主要開發語言，整合 OpenCV、Pygame、PyTorch、YOLOv7、BoT-SORT、ByteTrack 與 torchreid OSNet x0\_5 等元件，並以 OpenCV 與 Pygame 建立即時操作介面，提供即時框線顯示、畫面更新率呈現與滑鼠點選目標功能。依據目前完成成果，本作品已整合人物偵測、以 BoT-SORT 為主的追蹤器、重識別回復、非同步推論管線、雲台控制介面及視覺化操作等核心功能；其中 ByteTrack 保留作為原始對照基線。惟 TensorRT 加速、實際雲台參數調校與進階跨場景重識別仍有待後續驗證。整體而言，本作品已具備完整原型與應用發展潛力，能作為智慧監控與自動導引系統之實作基礎。

##### (一) 結論

本作品完成即時人物偵測、追蹤、重識別與雲台控制之整合設計，證明以模組化方式串接深度學習模型與控制迴路，可有效提升目標鎖定與追蹤之穩定性與即時反應能力。

##### (二) 未來發展

未來可導入 TensorRT 以提升高解析度場景之推論效率，並進一步完成真實雲台硬體之閉迴路控制測試與參數最佳化，同時擴充跨場景與跨鏡頭之重識別能力，以提升作品整體實用性與部署價值。本作品透過非同步推論架構、動態重識別機制與雲台閉迴路控制之整合，展現影像辨識與機構控制跨域結合之可行性，並可依不同應用需求替換偵測模型、追蹤策略或控制器，具備良好的模組化與擴充能力。

##### (三) 應用價值

本作品可延伸應用於智慧監控、自動巡檢、互動式展示與視覺導引等情境，並

可作為影像追蹤系統、機器視覺控制及嵌入式智慧裝置整合之教學與研究平台。

#### (四) 參考文獻說明

本報告所列文獻為作品實作過程中所參考之核心技術資料，依實際引用順序編排，格式參照 IEEE 標準撰寫。

#### 參考文獻

- [1] N. Aharon, R. Orfaig, and B.-Z. Bobrovsky, "BoT-SORT: Robust Associations Multi-Pedestrian Tracking," arXiv preprint arXiv:2206.14651, 2022.
- [2] Y. Zhang, P. Sun, Y. Jiang, D. Yu, F. Weng, Z. Yuan, P. Luo, W. Liu, and X. Wang, "ByteTrack: Multi-Object Tracking by Associating Every Detection Box," in Proceedings of the European Conference on Computer Vision, 2022.
- [3] K. Zhou and T. Xiang, "Torchreid: A Library for Deep Learning Person Re-Identification in PyTorch," 2019.
- [4] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors," arXiv preprint arXiv:2207.02696, 2022.
- [5] K. J. Åström and T. Hägglund, PID Controllers: Theory, Design, and Tuning, 2nd ed. Research Triangle Park, NC, USA: Instrument Society of America, 1995.