

提升誘捕籠剪耳辨識之可靠度：從耳朵計數到結構化判定條件

朝陽科技大學 [學號] [姓名]

[組員 2 / 學校 / 學號 / 姓名]

[組員 3]

[組員 4]

摘要

TNR(捕捉—絕育—回置)以剪除單側耳尖標記已完成絕育之個體。本團隊已建置一套於微控制器上執行剪耳辨識、並自主決定是否關閉閘門之誘捕籠原型。現行判定邏輯以單一模型同時偵測貓臉、剪耳與未剪耳三類；狀態機於偵測到貓臉後進入觀察狀態，於觀察窗內累加剪耳與未剪耳之偵測框數量，窗期結束時依兩者之比例決定是否捕獲。此設計於實際部署中顯現三項限制：其一，累加單位為偵測框個數而非影像幀數，同一姿態被重複計入，且流程不含結構性檢核；其二，模型以單階段訓練取得，未經場域資料之分階段微調，對角度與光線變化之適應能力不足；其三，剪耳正樣本數量結構性不足，直接影響可達到之辨識精度。本專題擬針對上述三點進行改良：以兩階段遷移學習提升特徵品質，以訓練時資料增強補償樣本不足，並以五項結構化檢查條件取代單純之數量比對，使系統得以區辨「影像不足以判定」與「影像足夠但未觀察到剪耳」兩種情形。預期成果為提高有效證據幀之比例，並使每一次判定具備可追溯之依據。本專題亦預期一種負面結果：若改良後仍無法取得足夠證據，則可據以確認瓶頸不在辨識模型，而在拍攝配置，此結論本身亦具價值。

1. 前言

研究動機。臺灣自 2017 年停止以安樂死作為族群控制手段後，收容所容量即成為主要限制[1]，降低繁殖率因而成為主要管理途徑。TNR 為現有文獻中研究最完整、對族群規模影響最明確之介入方式[2]，其成效取決於能否捕獲未絕育個體，而現場唯一可用之辨識依據為耳尖是否被剪除[3]。傳統誘捕籠僅具備觸發關門之機構，凡進入籠內之動物均會被捕獲。於持續實施 TNR 之區域，已絕育個體之比例隨時間累積；以本團隊之實測場域為例，約八成個體已完成絕育。在此情形下多數捕獲為重複捕獲，除耗費人力外，亦使動物承受不必要之緊迫；且因閘門以鏈條驅動、關閉時聲響明顯，鄰近貓隻可能因此習得迴避行為。

問題定義。現行判定流程如下：單一 YOLO 模型於每一幀同時輸出貓臉、剪耳、未剪耳三類偵測框；狀態機於任一幀偵測到貓臉時即進入觀察狀態，並將計數器歸零；其後於觀察窗內逐幀將剪耳與未剪耳之框數分別累加；窗期結束時，若未剪耳之累計框數超過

剪耳之一定倍數且本身達最低值，即判定未絕育並關閉閘門，否則放行。此流程在影像清晰時可正常運作，但於實際場域中顯現三項限制。第一，累加之單位為偵測框個數，而非影像幀數。同一隻動物於連續多幀維持相同姿態時，該姿態被重複計入，等同於同一項證據投票多次；若該姿態本身即為誤判，誤差隨之放大。第二，流程中不含任何結構性檢核：不檢查貓臉是否恰為一隻、耳朵是否恰為兩隻、耳朵之偵測框是否確實位於貓臉範圍內，亦不排除「同一個體兩耳皆被判為剪耳」此種依 TNR 慣例不可能發生之結果。側臉、歪頭或同一耳朵被重複框選時，數量條件仍可能成立，系統因而在證據不足下作成判定，且事後無從得知該次判定看到了什麼。第三，模型以單階段訓練取得，未經場域資料之分階段微調；加以 TNR 僅剪除單側耳尖，同一畫面必然同時出現一隻剪耳與一隻未剪耳，剪耳樣本結構性地少於未剪耳，而剪耳又是判定已絕育之唯一依據，遂成整體瓶頸。

現有方法與其不足。既有智慧誘捕籠之研究著眼於物種層級之區分：Besala 等人所提出之 AI 箱型陷阱，判斷畫面中之動物是否屬於目標物種[5]。本專題之任務與之不同——待區分之兩類為同一物種之兩種狀態，其差異僅在數公分之耳廓輪廓。運算位置配置（邊緣或雲端）已有延遲與能耗方面之研究[4]，惟該類文獻多假設偵測器之輸出可直接據以行動，未處理「裝置證據不足時應如何反應」。就本任務而言，僅提升平均精度並不足夠，尚須判斷單一幀是否具備作成判定之資格。

本專題之範圍。硬體、軟體與門控機構已完成建置並於場域運作，不在本專題範圍內。本專題只處理辨識模型之訓練方式，以及由偵測結果推導至捕獲動作之判定條件；兩者皆以既有裝置與既有資料改良，不新增硬體。上述限制均經現行程式碼與訓練設定檔比對確認。

預期採用之方法。針對前述三項限制，本專題擬以兩階段遷移學習取代單階段訓練、於訓練過程中施加資料增強，並以五項結構化檢查條件取代數量比對，使不符合結構條件之影像幀被標示為「不可用」而非逕行納入判定。細節見第 2 節。

2. 方法

2.1. 兩階段遷移學習。第一階段自 COCO 預訓練權

重出發，以通用貓隻影像（約1,700張原始影像）學習貓臉與耳朵之一般特徵；第二階段以第一階段之最佳權重為起點，以誘捕籠內實際拍攝之場域影像（約600張原始影像）微調，使模型適應部署環境之角度、光線與距離。場域資料不足以單獨訓練出穩定之特徵表示，通用資料雖多卻不含籠內拍攝條件，故須分階段使用。模型採 YOLOv7-tiny，輸入 640×640，以符合裝置端神經網路加速單元之限制。

2.2. 資料增強。資料集本身不預先增強，改於訓練過程中即時施加，以避免重複樣本佔用資料量。擬採用之項目包含 Mosaic 四張拼接（強度 1.0）、水平翻轉（0.5）、HSV 色彩擾動（H0.015 / S0.7 / V0.4）、平移與縮放（0.1 / 0.5）與少量 MixUp 及 Paste-in（各 0.05）。垂直翻轉與旋轉不予採用，因其不符合籠內之實際拍攝姿態。本專題將以增強前後之逐類別精度比較，檢驗增強對剪耳類別是否確有助益。

2.3. 結構化檢查條件。本專題擬以五項條件取代現行之數量比對，每一幀須依序通過方可成為有效證據：① 貓臉數量恰為 1；② 耳朵數量恰為 2；③ 兩耳不可皆判為剪耳；④ 兩耳之偵測框均須與貓臉框重疊；⑤ 三類別之信心值各自達門檻。其中條件 ③ 依據 TNR 僅剪除單側耳尖之慣例——兩耳皆為剪耳之情形實際上不會發生，出現即表示分類錯誤。條件 ①、②、④ 檢核影像結構是否合理，與門檻設定無關；此一分離使系統得以區辨兩種不同性質之失敗，並使每次判定之失敗原因可被記錄統計。同時，證據之累加單位由「偵測框個數」改為「通過檢查之有效幀數」，並設最短計數間隔，使同一姿態不致被重複計入。

2.4. 證據累積。通過結構檢查之有效幀，依剪耳信心值分入兩條並行軌道：高信心軌道（信心 ≥ 75）累積 2 幀即判定已絕育，低信心軌道（信心 ≥ 65）則需 3 幀。兩軌並行使系統於影像清晰時可提早完成判定，於影像品質普通時仍保有判定能力。有效幀之計數間隔至少 1 秒，避免同一姿態之連續幀重複累積。上述門檻與幀數皆集中定義於設定檔，以便後續調整。

2.5. 評估方式。本專題將逐幀記錄每一次觀察窗內之推論結果、各項結構條件之通過與否，以及當時生效之門檻值。據此可計算三項指標：有效證據幀之比例、各結構條件之拒絕分布，以及逐類別之召回率。第二項指標尤為關鍵：若拒絕主要來自信心門檻，則問題在模型精度；若主要來自結構條件，則問題在拍攝配置。兩者之改善方向完全不同。

2.6. 部署限制。所有改良皆須能於既有之 RTL8735B 微控制器上執行。模型須可量化為裝置端之執行格式，判定邏輯須能以既有韌體結構實作。任何無法滿足此條件之方案，即使離線評估表現較佳，亦不納入。

3. 預期結果

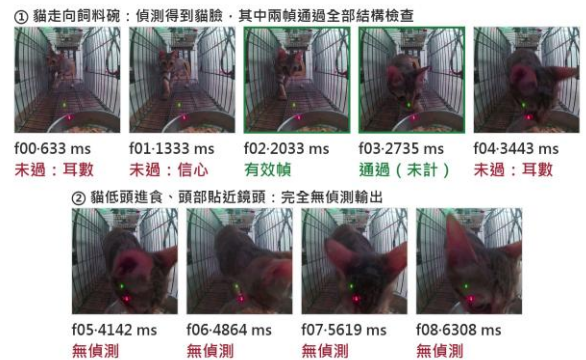


圖 1. 一次觀察窗內之連續影像（九幀，時間為自觀察起算之毫秒）。現行做法無法區分「影像不足以判定」與「影像足夠但未觀察到剪耳」；本專題預期以結構化檢查將前者標示出來，並記錄未通過之條件。

項目	現行做法	本專題目標
判定依據	耳朵框數比例	五項結構檢查＋雙軌累積
累加單位	偵測框個數	有效幀數（含節流）
訓練方式	單階段	兩階段遷移學習
不可用影像	無法辨識	標示並記錄拒絕原因

表 1. 現行做法與本專題目標之對照。

評估將以逐類別召回率呈現，而非整體正確率：在已絕育個體約佔八成之場域中，「一律判為已絕育」即可得到約八成之正確率，該數字不具參考價值。本專題預期之主要產出有三：有效證據幀比例之提升幅度、各結構條件之拒絕分布，以及剪耳類別在兩階段訓練與資料增強前後之精度變化。須說明的是，負面結果仍屬可能：若剪耳樣本之數量限制無法以演算法手段補償，則改良後之有效證據比例可能僅小幅提升。即使如此，拒絕分布仍可指出瓶頸所在；能將此二成因明確分離並量化，即為本專題之實質貢獻。

參考文獻

- [1] Yan, T.-Y., and Teng, K. T.-y. Trends in animal shelter management, adoption, and animal death in Taiwan from 2012 to 2020. *Animals*, 13:1451, 2023.
- [2] Smith, L. M., Hartmann, S., Munteanu, A. M., Dalla Villa, P., Quinell, R. J., and Collins, L. M. The effectiveness of dog population management: A systematic review. *Animals*, 9:1020, 2019.
- [3] Agriculture Knowledge Portal, Taiwan. Introduction to TNR. 2020.
- [4] Koubaa, A., Ammar, A., Alahdab, M., Kanhouh, A., and Azar, A. T. DeepBrain: Experimental evaluation of cloud-based computation offloading and edge computing in the Internet-of-Drones for deep learning applications. *Sensors*, 20:5240, 2020.
- [5] Besala, F. I., Niimoto, R., Lee, J. H., and Okamoto, S. Development of AI-based smart box trap system for capturing a harmful wild boar. *ROBOMECH Journal*, 12:4, 2025.