

智慧誘捕籠：剪耳辨識

邱識瑋
朝陽科技大學
11530601

吳翊銘
朝陽科技大學
11430624

摘要

本專題用影像辨識判斷流浪貓是否剪過耳。剪耳是 TNR（捕捉、絕育、回置）中，現場唯一看得出來的結紮記號。我們的智慧誘捕籠把 YOLOv7-tiny 放在籠內微控制器上，沒偵測到剪耳就關門。第一版在貓店運作四個月、509 次判定，86.1% 沒有做出判斷，證據不足時直接放走；以 111 張照片離線重評，剪耳只認出 19.6%。本專題提出第二版處理兩個卡點：剪耳樣本太少，以補樣本、重複抽樣和 focal loss 處理；照片品質沒有把關，改為先過五項檢查、多張累積才下結論，並讓「無法判定」成為明確的結果，不確定就不下結論。本學期做出第二版，與第一版在同一批資料上對照。

1. 前言



圖 1：完整耳與剪耳，只差在耳尖。

研究動機。臺灣 2017 年停止以安樂死控制族群數量後，收容所容量成為主要限制 [1]，降低繁殖率成為主要的管理方式。TNR 是研究最多、對族群數量影響最明確的做法 [2]，但成敗在於能不能抓到未結紮的貓，而現場唯一的線索是耳尖有沒有被剪（圖 1）[3]。傳統誘捕籠走進去就關門，所以在長期做 TNR 的地區，多數捕獲其實是重複抓到已結紮的貓。

這件事難在哪。第一，差別太細：只差在一邊耳尖那幾公分的形狀。第二，數量天生不平衡：TNR 只剪一邊，已結紮的貓也有一耳完整，所以剪耳永遠是少數，偏偏它是判斷結紮的唯一依據。第三，單張照片不可靠：

貓的姿勢和距離一直在變，多數照片不足以下判斷，第一版就卡在這裡（圖 2）。現有的智慧誘捕籠分的是長相差很多的物種 [4]，沒有處理照片不夠判斷的情況。



圖 2：第一版的判斷規則與四個月的實測成績。證據不足時直接放走，因此多數進籠沒有結論。

第一版做了什麼。第一版（2025 年 12 月至 2026 年 4 月）每次進籠觀察 7 秒、約 9 張照片，每張用同一個門檻 40 分，不分類別；剪耳出現 2 次且平均 40 分以上判已結紮、放行；耳朵偵測不足 5 次就重試，最多 3 次（圖 2）。

第一版卡在哪（表 1）。兩台裝置在貓店運作四個月、509 次判定，86.1% 沒有做出判斷；重試用完仍然不足時直接放走，等於沒抓。以 111 張照片離線重評，剪耳召回率只有 19.6%。原因有兩個：剪耳樣本天生偏少；判斷規則只數次數，不管每張照片夠不夠格，模糊的和清楚的一樣算。本專題以現有裝置做出第二版，從資料、訓練和判斷規則同時處理（圖 3）；就算方法沒用，也要找出卡在哪裡。

項目	結果
現場判定次數	509
沒有做出判斷	86.1%
證據不足時的動作	重試 3 次後放走
離線重評照片	111 張
剪耳召回率	19.6%
每張照片的品質檢查	無

表 1：第一版實測結果（兩台裝置，貓店，2025 年 12 月至 2026 年 4 月）。

2. 方法

2.1. 模型與裝置

YOLOv7-tiny [5]，輸入 640x640，辨識貓臉、剪耳、完整耳三類，從 COCO 預訓練權重開始訓練。訓練後

量化成 int16 (改用整數權重, 裝置才跑得動), 在籠內的 RTL8735B 微控制器上執行, 每張約 0.7 秒, 不連伺服器。



圖 3：第二版的方法流程。上半在電腦上訓練，下半在裝置上逐張辨識並累積判斷；綠色為第二版新增，灰色為第一版已有。

2.2. 資料與切分

第一階段用網路蒐集的一般貓照 1,729 張 (含剪耳 793 張; 剪耳約 800、完整耳約 2,700 個標註, 1:3.3); 第二階段用籠內照 606 張 (含剪耳 144 張; 約 145:490, 1:3.4), 其中 291 張是紅外線觸發但沒有貓的空籠, 當作負樣本, 教模型不要對背景亂框。兩種資料都依 70/15/15 切成訓練、驗證、測試組, 按「含剪耳/只有完整耳/空籠」分層, 亂數種子固定。增強只用在訓練組, 挑模型只看驗證組, 測試組留到最後只算一次。

2.3. 遷移學習與資料增強

第一階段用一般貓照學貓臉和耳朵的樣子; 第二階段從第一階段最好的權重出發, 用籠內照微調 (籠內照太少, 單獨訓練學不穩), 適應籠內的角度、光線和距離。訓練時即時增強: Mosaic (四張拼一張)、MixUp (兩張疊合)、左右翻轉、HSV 擾動、平移、縮放和 Paste-in (把物件貼到別張), 但不做上下翻轉和旋轉, 因為籠內的貓不會倒著出現。

2.4. 處理剪耳太少

籠內含剪耳的照片只有 144 張, 所以本專題 (1) 補資料: 第二版韌體存下每次進籠的每一張照片, 用來補標剪耳; (2) 重複抽樣: 含剪耳的照片每輪多抽幾次 [6]; (3) 用 focal loss [7] 降低容易樣本的權重, 專注在難認的剪耳; (4) 挑模型改看剪耳召回率, 不看三類平均, 因為判斷結紮只靠剪耳。

2.5. 多張累積

每次進籠觀察 6 秒、約 9 張。每張須過五項檢查才算有效張: 貓臉恰 1 個、耳朵恰 2 個、不可兩耳皆剪耳 (TNR 只剪一邊)、兩耳都與貓臉重疊、信心達門檻 (貓臉 65、耳朵 50, 滿分 100)。有效張至少間隔 1 秒, 避免同一姿勢重複計算。剪耳 75 分以上累積 2 張, 或 65 分以上累積 3 張, 即判已結紮。第二版把「無法判定」變成明確的第三種結果: 沒有任何有效張就不下結論、不關門, 並記下是哪一項檢查擋掉的; 第一版遇

到證據不足是重試 3 次後直接放走, 事後查不出卡在這裡。這是拒絕判斷 (reject option) 的概念 [8]: 錯抓已結紮的貓無法挽回, 漏抓的貓下次還會再來; 兩種錯的代價不對稱, 理性代理人應讓預期損失最小 [9]。

2.6. 怎麼驗證

模型層看測試組上每一類的 AP 與召回率, 重點是剪耳。判斷層用韌體記下的每張分數與檢查結果, 離線回放第一版與第二版的規則在同一批進籠上的表現。不看整體正確率: 現場約八成已結紮, 一律回答「已結紮」就有八成正確率。

3. 預期結果

本專題將交出: (1) 各種剪耳補強方法在測試組上的剪耳召回率, 與第一版對照; (2) 新舊規則在同一批進籠上的回放結果, 重點是「沒有做出判斷」的比例能不能從 86.1% 降下來; (3) 五項檢查各擋掉多少張, 用來分辨問題出在模型還是拍攝。可預見的限制: 被擋下的照片多半是臉或耳朵不完整, 光改模型可能救不回來。實驗結果與錯誤分析留待後續作業報告。

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., Quinnell, 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] 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.
- [5] Wang, C.-Y., Bochkovskiy, A., and Liao, H.-Y. M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In *CVPR*, 2023.
- [6] Gupta, A., Dollár, P., and Girshick, R. LVIS: A dataset for large vocabulary instance segmentation. In *CVPR*, 2019.
- [7] Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P. Focal loss for dense object detection. In *ICCV*, 2017.
- [8] Chow, C. K. On optimum recognition error and reject tradeoff. *IEEE Transactions on Information Theory*, 16(1):41 – 46, 1970.
- [9] Russell, S., and Norvig, P. *Artificial Intelligence: A Modern Approach*. Pearson, 4th edition, 2020.