

# 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朵。剪耳是 TNR（捕捉、絕育、回置）中，也就是現場唯一看得出來的是已經結紮的記號。我們的智慧誘捕籠把 YOLOv7-tiny 放在籠內的微控制器上，如果沒有偵測到剪耳就會關門，去做結紮的動作。目前的系統在貓咪收容所運作約四個月，累計 509 次判定，但是有 86.1% 沒有辦法做出有效判斷，後續以 111 張照片重新評估後發現，剪耳只能判定出 19.6%。

整理實測紀錄後，我們找出三個原因：剪耳樣本太少、把單張照片當成完整資訊，在證據不足時沒有明確的處理。所以我們參考本了機器學習與智慧代理人觀念，規劃三個改善方向：補足剪耳樣本並用遷移學習、累積多張照片的證據再判斷是否結紮剪耳，透過這段期間實驗與驗證結果。

## 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，目前的系統就卡在這裡（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. 目前的判斷方式與四個月的實測成績。證據不足時重試後放走，因此多數進籠沒有結論。

**目前遇到的問題（表 1）。**目前的系統每次進籠觀察約 7 秒、拍約 9 張，每張用同一個門檻，數剪耳出現幾次來決定放行或關門；證據不夠就重試，重試用完就放走（圖 2）。兩台裝置在貓店運作四個月、509 次判定，86.1% 沒有做出判斷，等於沒抓；以 111 張照片離線重評，剪耳召回率只有 19.6%。

| 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. 目前系統的實測結果（兩台裝置，貓店，2025 年 12 月至 2026 年 4 月）。

**我們找到的原因。**整理實測紀錄後，我們認為問題有三個來源。第一，剪耳樣本太少，模型學得不夠。第二，把單張照片當成完整資訊：用課程的話說，目前的系統是簡單反射型代理人（simple reflex agent），看到一張照片就照固定規則反應；但籠內是部分可觀察（partially observable）的環境，單張照片看不到完整的狀態 [5]。第三，證據不足時沒有明確的處理：模糊的照片和清楚的照片一樣算，卡住了只能重試再放走，事後也查不出卡在哪裡。

**本專題的規劃。**透過這門課與目前的進度，我們針對

三個原因各規劃一個改善方向（第2節、圖3）。這份提案先定方向，實際做法會邊做邊調整；就算某個方向沒有效果，也要找出卡在哪裡。

## 2. 方法

### 2.1. 模型與裝置

偵測模型是YOLOv7-tiny [6]，辨識貓臉、剪耳、完整耳三類；訓練後壓縮成整數格式（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。以下三個構想都在這個條件下進行，不更換硬體。



圖3. 規劃的改善方向。上半在電腦上訓練，下半在裝置上判斷；綠色是本專題規劃的部分，構想一、二、三分別對應第2.2~2.4節。

### 2.2. 構想一：補足剪耳樣本

這一點對應課程的機器學習：模型學得到什麼，取決於給它的資料。我們手上有網路蒐集的一般貓照約1,700張與籠內照數百張，剪耳和完整耳的標註約1:3，剪耳明顯是少數。規劃的方向是：先用一般貓照學貓臉和耳朵的樣子，再用籠內照微調（遷移學習），彌補籠內照太少的問題；持續補標籠內拍到的剪耳照片，空籠照片當作負樣本；並試用重複抽樣 [7] 或 focal loss [8] 這類處理類別不平衡的方法。資料會分成訓練、驗證、測試三組，挑模型只看驗證組，避免過度擬合（overfitting，只記住訓練資料、換一批就不準） [5]。

### 2.3. 構想二：累積多張照片再判斷

這一點對應課程的代理人設計與不確定性推理。面對部分可觀察的環境，代理人應該保留內部狀態，把一連串的感知累積起來再決定，而不是只看眼前這一張（model-based agent） [5]。我們規劃把判斷改成兩步：每張照片先做品質把關（例如臉和耳朵是否完整、信心是否足夠），不合格的不算；合格的照片再累積剪耳證據，證據夠了才下結論。把關與累積的條件要實際試過才定得下來。

### 2.4. 構想三：讓系統會說「不知道」

這一點對應課程的效用型代理人（utility-based agent）：理性的代理人應選擇預期效用最大的動作 [5]。這個任務裡兩種錯的代價不對稱：錯抓已結紮的貓無法挽回，

漏抓的貓下次還會再來。所以證據不足時，比起硬猜，更好的動作是明確回答「無法判定」、不關門，這就是拒絕判斷（reject option）的概念 [9]。每次無法判定也記下是哪一步擋掉的，當作改善模型和規則的回饋，類似學習型代理人中「評論者」（critic）的角色 [5]。

### 2.5. 怎麼驗證

依課程的 PEAS 架構，先把效能指標（performance measure）定清楚 [5]：看剪耳這一類的召回率，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，這個數字沒有意義。判斷規則則用記下的照片在同一批進籠上離線重播，比較新舊做法。

## 3. 預期結果

我們預期：(1) 補足樣本並用遷移學習後，剪耳召回率能高於目前的19.6%；(2) 「沒有做出判斷」的86.1%能被拆開，分成「照片真的不夠」和「其實可以判」兩種，並知道每次卡在哪一步；(3) 證據不足時系統會明確標成「無法判定」並記下原因，而不是重試後放走、查不出問題。也可能遇到限制：如果照片本身就拍不到臉或耳朵，改模型也救不回來，但記下的原因能幫我們分辨是模型還是拍攝的問題。實際數值、調整過程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] Russell, S., and Norvig, P. *Artificial Intelligence: A Modern Approach*. Pearson, 4th edition, 2020.
- [6] 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.
- [7] Gupta, A., Dollár, P., and Girshick, R. LVIS: A dataset for large vocabulary instance segmentation. In *CVPR*, 2019.
- [8] Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P. Focal loss for dense object detection. In *ICCV*, 2017.
- [9] Chow, C. K. On optimum recognition error and reject tradeoff. *IEEE Transactions on Information Theory*, 16(1):41 – 46, 1970.