

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

本專題是用影像辨識的方式，來判斷一隻流浪貓有沒有被剪過耳朵。在 TNR（捕捉、絕育、回置）的做法中，剪耳就是現場唯一看得出來這隻貓已經結紮的記號。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，把貓抓起來帶去做結紮。目前的系統在貓咪收容所運作了大約四個月，累計 509 次判定，但是有 86.1% 沒有辦法做出有效的判斷；後續我們用 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，以及在證據不足時沒有明確的處理。所以我們參考了本課程的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有剪耳，以及證據不夠時讓系統回答「無法判定」、先不要關門，並且在這段期間透過實驗來驗證結果。

1. 前言



圖 1. 完整耳和剪耳的差別，只在耳尖的形狀。

研究動機。臺灣在 2017 年開始不再用安樂死的方式來控制流浪動物的數量，從那之後，收容所能容納多少動物就變成主要的限制 [1]，所以降低繁殖率就成為主要的管理方式。TNR 是被研究最多、而且對控制數量效果最明確的方法 [2]，但是它能不能成功，關鍵在於能不能抓到還沒結紮的貓，而現場唯一的線索，就是耳尖有沒有被剪掉（圖 1）[3]。不過傳統的誘捕籠只要貓走進去就會關門，所以在長期做 TNR 的地區，大部分

抓到的其實都是已經結紮過的貓。

這件事難在哪。這個問題難的地方有三個。第一是差別太細，因為剪耳和沒剪耳只差在一邊耳尖那幾公分的形狀。第二是數量本來就不平衡，因為 TNR 只剪一邊，就算是已經結紮的貓也還有一隻耳朵是完整的，所以剪耳的資料永遠比較少，但是它偏偏又是判斷有沒有結紮的唯一依據。第三是單張照片不可靠，因為貓在籠子裡的姿勢和距離一直在變，大部分的照片其實都不夠拿來判斷，目前的系統就是卡在這裡（圖 2）。另外，現有的智慧誘捕籠研究分辨的是長得差很多的物種 [4]，並沒有處理到照片不夠判斷的情況。

目前的判斷方式（2025 年 12 月～2026 年 4 月實測）



圖 2. 目前的判斷方式和四個月的實測成績。證據不夠時會重試，試完就把貓放走，所以大部分進籠都沒有結論。

目前遇到的問題。目前的系統在貓進籠之後，會觀察大約 7 秒、拍大約 9 張照片，每一張都用同一個門檻，然後數剪耳出現了幾次，再決定要放行還是關門；如果證據不夠就重試，試到次數用完就把貓放走（圖 2）。我們用兩台裝置在貓店運作了四個月，總共 509 次判定，但是其中有 86.1% 沒有做出判斷；另外用 111 張照片重新評估，剪耳也只能認出 19.6%。

我們找到的原因。整理實測紀錄之後，我們認為問題主要有三個來源。第一是剪耳的樣本太少，所以模型學得不夠好。第二是系統把單張照片當成完整的資訊，用這門課的說法，目前的系統是一個簡單反射型代理人（simple reflex agent），看到一張照片就照固定的規則反應；但是籠子裡其實是部分可觀察（partially observable）的環境，只看一張照片是看不到完整狀況的 [5]。第三是證據不足時沒有明確的處理，模糊和清楚的照片都被當成一樣來算，卡住的時候只能重試再放走，事後也查不出到底卡在哪裡。

本專題的規劃。因此，我們結合這門課學到的觀念和目前的進度，針對這三個原因各規劃了一個改善方向（第 2 節、圖 3）。實際的做法會一邊做一邊調整。

2. 方法

2.1. 模型與裝置

我們使用的偵測模型是 YOLOv7-tiny [6]，它會辨識貓臉、剪耳和完整耳三個類別。模型訓練完之後會先壓縮成整數的格式，這樣裝置才跑得動，然後放在籠子裡的 RTL8735B 微控制器上執行，每一張照片大約需要 0.7 秒，不需要連到伺服器。



圖 3. 我們規劃的改善方向。上半部是在電腦上訓練，下半部是在裝置上判斷；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

這個構想對應到課程中機器學習的部分，因為模型能學到什麼，取決於我們給它什麼樣的資料。目前我們有從網路蒐集的一般貓咪照片大約 1,700 張，以及籠子裡拍到的照片幾百張，剪耳和完整耳的標註比例大約是 1:3，剪耳明顯是少數。所以我們打算先用一般的貓咪照片讓模型學會貓臉和耳朵長什麼樣子，再用籠子裡的照片做微調，也就是遷移學習；同時持續把籠子裡拍到的剪耳照片補進來，並且試試看重複抽樣 [7] 或 focal loss [8] 這類處理資料不平衡的方法。資料會分成訓練、驗證、測試三組，挑模型的時候只看驗證組，避免過度擬合 (overfitting，也就是模型只記住訓練資料，換一批就不準) [5]。

2.3. 構想二：累積多張照片再判斷

這個構想對應到課程中代理人的設計。課程提到，在部分可觀察的環境裡，代理人應該要保留一個內部狀態，把一連串看到的東西累積起來再做決定，而不是只看眼前這一張 (model-based agent) [5]。所以我們打算把判斷改成兩個步驟：第一步是每一張照片先做品質把關，例如臉和耳朵是不是完整、信心分數夠不夠高，不合格的不算；第二步是把合格的照片累積起來，剪耳的證據夠了才下結論。至於條件要怎麼設定，要實際試過之後才能決定。

2.4. 構想三：讓系統會說「不知道」

這個構想對應到課程中的效用型代理人 (utility-based agent)，也就是理性的代理人應該要選擇預期效用最大的動作 [5]。在我們這個題目裡，兩種錯誤的代價不一樣：如果把已經結紮的貓抓起來，這是沒辦法挽回的；但是如果漏抓了一隻還沒結紮的貓，牠下次還是

會再來。所以在證據不夠的時候，與其硬猜一個答案，比較好的做法是讓系統回答「無法判定」，先不要關門，這就是拒絕判斷 (reject option) 的概念 [9]。另外，系統也會記下每一次是在哪一步被擋下來的，拿來當作改善的回饋，類似課程中學習型代理人裡「評論者」 (critic) 的角色 [5]。

2.5. 怎麼驗證

我們會依照課程中的 PEAS 架構，先把效能指標 (performance measure) 定清楚 [5]。我們主要看的是剪耳的召回率，而不是整體的正確率，因為現場大約有八成的貓已經結紮了，就算系統全部都回答「已結紮」，正確率也會有八成，所以這個數字沒有什麼意義。判斷規則的部分，則是用記錄下來的照片重新跑一次，比較新舊做法的差別。

3. 預期結果

我們預期可以達到三件事。第一，補足樣本並且使用遷移學習之後，剪耳的召回率能比目前的 19.6% 還要高。第二，目前「沒有做出判斷」的 86.1% 能被拆開來看，分成「照片真的不夠」和「其實可以判斷」兩種，並且知道每一次是卡在哪一步。第三，證據不夠的時候，系統會明確標記成「無法判定」並記下原因，而不是重試之後就把貓放走。不過如果照片本身就拍不到臉或耳朵，改善模型也沒辦法解決。實際的數值和錯誤分析，會在之後的作業中再報告。

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.