

從固定門檻到效用導向代理人：邊緣運算剪耳辨識誘捕籠之不確定性 感知捕獲決策

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) 計畫以剪除單側耳尖作為已絕育個體之標記。傳統誘捕籠無法區分剪耳與未剪耳個體，因此在已絕育比例約八成之場域中，多數捕獲屬徒勞，並可能使貓隻習得迴避誘捕籠。本團隊已部署一套邊緣運算誘捕籠，於微控制器等級之神經網路加速器上執行剪耳偵測，並自主決定是否關閉閘門。116 筆現場判定紀錄顯示，問題出在決策層而非感知層：69% 之判定在完全沒有有效證據幀的情況下作成，34 次捕獲中有 24 次 (71%) 並無任何支持證據，違反系統自身「不確定時不得捕獲」之原則。本專題以該原型為基準線，將捕獲決策重新形式化為「部分可觀測環境下之效用導向理性代理人」，以明確之信念更新與非對稱成本決策，取代人工調校之固定門檻。感知層視為既有條件，本專題之工作完全集中於決策層。由於每筆判定紀錄皆保存逐幀信心值與當時生效之門檻值，新策略可直接以既有資料進行離線回放開發與評估。預期成果為消除零證據捕獲、同時維持捕獲召回率，並產出可直接於現有微控制器上執行之精簡策略。

1. 前言

動機。自 2017 年臺灣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首要管理途徑。TNR 為其中研究最完整、且對族群規模影響最大之介入方式 [2]。其成效取決於能否捕獲未絕育個體，而未絕育個體唯一之辨識依據為耳尖未被剪除 [3]。傳統誘捕籠無從區分，凡進入者一律捕獲。本研究場域中約八成個體已完成絕育，因此多數捕獲為重複捕獲：耗費志工人力、造成動物緊迫，且由於鏈條驅動之鐵製閘門聲響甚大，可能使鄰近貓隻習得迴避行為，反而降低日後接觸真正目標個體之機會。

問題定義。本任務並非「偵測貓隻」，而是：在籠內固定相機所產生之一連串充滿雜訊且斷續的逐幀偵測結果之下，決定何時捕獲、何時放行、何時繼續觀察——且須在無網路保證之裝置上、於短暫觀察窗內、面對高度不對稱之成本完成。誤捕之代價昂貴且部分不可逆：動物緊迫、迴避誘捕籠、人力浪費；誤放之代價則低廉，因為同一個體仍會再度造訪。形式上，

此為部分可觀測、隨機、序列式任務環境中之序列決策問題，其正確目標為期望效用，而非逐幀分類準確率。

現有方法與其不足。既有智慧誘捕籠著眼於物種層級之區分；Besala 等人之 AI 箱型陷阱判斷畫面中之動物是否屬於目標物種 [5]。本任務性質不同：兩個類別為同一物種之兩種狀態，差異僅在數公分之耳廓輪廓，且具決定性之類別嚴重樣本不足。邊緣與雲端之運算位置配置已有延遲與能耗方面之研究 [4]，惟該文獻假設偵測器輸出可直接據以行動，並未處理裝置自身證據不足時應如何反應。

範圍。上述量測來自本團隊於研究場域實際運作之原型誘捕籠，該原型提供感知層——一個量化部署於裝置 NPU 上之剪耳偵測器——以及韌體、門控與逐幀證據記錄。本專題將感知層視為既有條件，處理其上一層：把不確定之偵測串流轉換為實體動作之策略。目前使用之決策邏輯為固定門檻規則，其診斷價值明確——該規則揭露約四分之三之結構檢查拒絕屬結構性因素而非門檻因素，指向相機安裝位置而非模型準確率——但作為決策機制則相當脆弱。其門檻為人工設定之常數，其累積器並無「證據累積至多少才算足夠」之概念，且「未觀察到剪耳」被逕行視為「未絕育」，此即 71% 之捕獲在零證據下作成之原因。以下提出之信念與效用形式化，目前並不存在於韌體之中。

提出之方法。我們以明確之信念與效用形式化取代固定門檻累積器。偵測器信心值先經校準轉換為概似度，再用以在觀察窗內更新對 {已絕育, 未絕育} 之後驗信念。每一步驟中，代理人依非對稱成本矩陣下之期望效用，在 CAPTURE (捕獲)、RELEASE (放行) 與 OBSERVE (繼續觀察) 之間選擇，其中 OBSERVE 帶有受實體觀察窗限制之時間成本。第三種結果 UNCERTAIN (不確定) 因而首次得以被表達。最終策略將回編為精簡之決策表，使微控制器韌體之結構維持不變。

2. 方法

2.1. 任務環境界定。我們先以 PEAS 架構界定此代理人 [6]。績效量度：捕獲與放行之期望成本，而非原始準確率。環境：籠內空間，一次一隻貓，其他貓隻可透過網格看見。致動器：閘門馬達與遠端通知。感測器：相機與兩組紅外線感測。此環境為部分可觀測、

隨機、序列式且單一代理人。此歸類決定可行解法之範圍，亦為逐幀分類器在結構上不足以勝任之原因。

2.2. 信心值校準。偵測器輸出之原始信心值並非機率。我們於保留集上擬合單參數之溫度縮放 [7]，使剪耳類別之信心值成為可用之概似度，並以期望校準誤差與可靠度圖評估校準品質。此步驟為必要，因為下游之信念更新會將這些數值相乘，校準偏差將直接傳遞為錯誤之動作。

2.3. 觀察窗內之信念更新。觀察窗被視為一連串觀測。未通過結構檢查之幀被視為「遺漏」而非負面證據；僅此一項改動即可消除「沒有證據等同於證據不存在」之缺陷。有效幀則用以更新後驗。我們將比較獨立概似度更新，與一個雙狀態隱藏馬可夫模型之形式化——後者反映單次造訪內底層狀態固定不變、但可觀測性隨時間變化此一事實。

2.4. 效用導向之動作選擇。我們定義一個成本矩陣，其非對稱項次反映前述場域現實，並選擇期望效用最大之動作。由於 OBSERVE 可行但有代價，且觀察窗上限為 6 秒（依實測 703 毫秒之幀間隔中位數計算約為九幀），此為最佳停止問題，以剩餘觀察窗之反向歸納求解。

2.5. 以回放進行離線策略評估。每筆判定紀錄皆保存逐幀信心值、各項條件之拒絕計數與當時生效之門檻值，因此任何候選策略皆可於已記錄之 116 筆判定與 1,125 幀上評分，無需重新至現場採集資料。此特性使單一學期內之快速迭代成為可能，亦為本專題可行性之主要依據。

2.6. 部署限制。最終策略將量化為以（累積信念，剩餘幀數）為索引之查找表，以符合既有 RTL8735B 韌體。任何無法如此編譯之策略即予排除。此限制與本團隊先前之架構選型考量一致：RT-DETR-L 雖僅損失 0.087 之 mAP，量化後卻完全失去決策層覆蓋率。

3. 預期結果



圖 1. 一次已記錄之觀察窗（九幀，0.6 – 6.3 秒）。其中僅 f02 成為有效證據幀，其餘皆未通過結構檢查。於現行規則下，此觀察窗將在幾乎沒有證據之情況下以捕獲作結；所提策略預期將其歸為 UNCERTAIN。

指標	現行固定門檻規則	本專題目標
零證據捕獲	24 / 34 (71%)	0
零證據判定	69%	列為 UNCERTAIN
未絕育召回率	尚無法計算	逐類別 + 信賴區間
裝置端策略	人工設定常數	查找表

表 1. 相對於現行規則之目標值。

我們將回報逐類別召回率而非整體準確率：在已絕育個體佔八成之場域中，一個「一律判為已絕育」之退化策略即可達到約 80% 之準確率。主要產出為一條決策成本曲線，於一系列成本比值下比較所提策略與現

行規則。負面結果仍有可能：若剪耳類別經校準後之信心值本身過弱，任何決策策略皆無法補救，屆時結論將轉為一項量化陳述——瓶頸在於感知層而非決策邏輯。能夠乾淨地分離此二成因，本身即為一項結果。

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] Koubaa, A., Ammar, A., Alahdab, M., Kanhouch, 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.
- [6] Russell, S., and Norvig, P. *Artificial Intelligence: A Modern Approach*. Pearson, 4th edition, 2020.
- [7] Guo, C., Pleiss, G., Sun, Y., and Weinberger, K. Q. On calibration of modern neural networks. In *ICML*, 2017.