

YOLO-Puppy：寵物犬骨架偵測及情緒分析系統

YOLO-Puppy: Pet Dog Body Keypoint Detection and Emotion Analysis System

研究生：資訊工程系 謝家奕

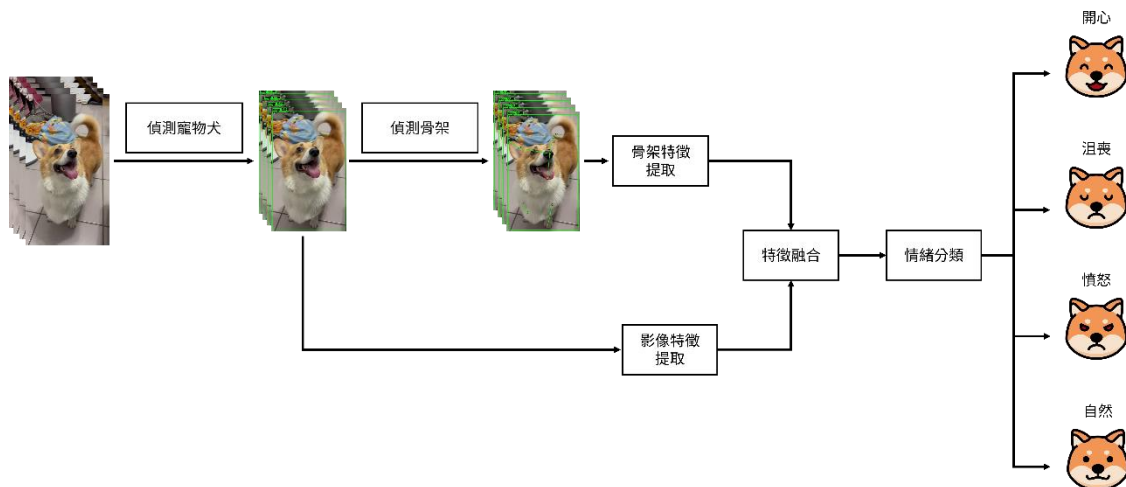
指導教授：資訊工程系 張傳育特聘教授

摘要

為了讓飼主們可以更加了解自家的寵物，本論文提出一種寵物犬情緒分析系統，以應對愈加普及的人與寵物犬相處的需求。本研究首先使用改良後的 YOLOv8n 模型進行寵物犬的骨架偵測，隨後將骨架資訊以及影片分別輸入進 CTR-GCN 模型和 MobileNetV3 進行特徵提取，其中 MobileNetV3 提取出來的特徵會經由 Bi-LSTM 進行時序處理，最終透過 Co-attention 模組將兩者的特徵融合並輸入到分類器以得出寵物犬的情緒。實驗結果表明，在寵物犬之偵測任務方面，模型的結果 mAP 達到了 99.4%，同時模型也在寵物犬節點偵測任務中獲得 94.3% 之 PCK 準確率，本論文提出之寵物犬情緒分析系統也在寵物犬情緒分類任務上準確率為 87.1%。上述的實驗結果表明了本論文提出之寵物犬分析系統的可行性。此系統的應用有望使飼主更加了解寵物犬的情感需求，進而採取相應的措施，促進與寵物犬之間更良好的互動。這不僅有益於飼主本身的心理健康，同時也有助於提升寵物犬的生活品質，實現雙贏局面。

研究方法

在本論文的實驗中，我們會先進行寵物犬及其身體節點偵測，並將偵測到的寵物犬骨架資訊（身體節點座標）以及原影片分別輸入到不同的特徵提取模型進行特徵提取。隨後，本研究融合兩者的特徵並輸入到後續的寵物犬情緒分類器中，最終輸出寵物犬的情緒結果。系統流程圖如下所示。



寵物犬及其身體節點偵測方面，本論文使用了 YOLOv8n 物件檢測模型作為主框架，對其架構進行了調整並用於偵測寵物犬以及其身體節點。寵物犬影片會設定成 15FPS，並以每 5 秒為一個單位進行輸入。在寵物犬及其身體節點偵測階段，模型會逐幀進行寵物犬的偵測。若該幀沒有偵測到寵物犬，則會跳過該幀改而輸入下一幀的影像進行寵物犬偵測；反之，若模型在該幀有偵測到寵物犬，則會對該寵物犬進行身體節點偵測，標記出耳朵、眼睛等 24 個不同的身體節點。

在獲得寵物犬的骨架資訊之後，本論文會將其輸入到 CTR-GCN 模型（Channel-wise Topology Refinement Graph Convolutional Network）進行骨架特徵的提取。同時，原影片會輸入到 MobileNetV3 模型中進行特徵提取並且輸入到 Bi-LSTM（Bidirectional Long Short-Term Memory）模型進行時序處理。本論文透過協同注意力模組（Co-Attention Module）將兩者的特徵提取結果進行融合，並輸入到包含 SoftMax 損失函數之情緒分類器中進行情緒分類，最終獲得寵物犬的情緒分析結果如開心、憤怒、沮喪和自然。

實驗結果

寵物犬與其身體節點偵測模型比較結果

偵測目標	DeepLabCut ^[1]	Ours
寵物犬 (mAP)	97.6%	99.4%
寵物犬節點 (PCK)	91.3%	94.3%

[1]: A. Mathis, P. Mamidanna, K.M. Cury et al., "DeepLabCut: markerless pose estimation of user-defined body parts with deep learning," Nature Neuroscience, vol. 21, pp. 1281–1289, 2018.

不同寵物犬情緒分析研究之模型比較結果

K-Fold	Kim Ferres ^[2]	Yan Mao ^[3]	Ours
1	69.4%	84.1%	84.7%
2	71.8%	86.5%	89.2%
3	70.7%	84.3%	87.5%
4	71.2%	85.7%	87.7%
5	70.5%	84.5%	86.6%
Mean	70.7%	85.0%	87.1%

[2]: K. Ferres, T. Schloesser, and P. A. Gloor, "Predicting Dog Emotions Based on Posture Analysis Using DeepLabCut," Future Internet, vol. 14, no. 4, pp. 97, 2022.

[3]: Y. Mao, Y. Liu, "Pet dog facial expression recognition based on convolutional neural network and improved whale optimization algorithm," Scientific Reports, vol. 13, pp. 3314, 2023.

結論

本研究對 YOLOv8n 模型進行調整並應用於寵物犬及其身體節點的偵測任務上，隨後將獲得的骨架資訊以及影片分別輸入進 CTR-GCN 模型和 MobileNetV3 進行特徵提取，最終透過 Co-attention 模組將兩者的特徵融合以得出寵物犬的情緒。實驗結果表明，在寵物犬之偵測任務方面，模型的結果 mAP 達到了 99.4%，同時模型也在寵物犬節點偵測任務中獲得 94.3% 之 PCK 準確率，本論文提出之寵物犬情緒分析系統也在寵物犬情緒分類任務上的準確率為 87.1%。實驗結果表明本研究提出的方法為寵物犬相關領域帶來了更加準確的自動化工具，其可以應用於寵物犬醫療保險、智能監控系統等實際應用中，進而提升寵物的生活品質以及促進人與寵物之間的互動。未來本團隊希望可以將這些技術擴展到其他動物種類的姿態辨識與情緒分類任務上，實現人類與寵物可以互相理解的理想。