

透過深度學習理解狗的情緒：基於 CNN 的分類框架

維舍什·坦瓦爾
印度旁遮普省奇特卡拉大學工程技術學院

vishesh.tanwar@chitkara.edu.in

摘要—我們的研究文章重點在於開發基於深度學習的捲積神經網路 (CNN) 模型，將狗的情緒分為四類：悲傷、快樂、放鬆和憤怒。強調了了解犬類情緒的重要性，強調需要準確的分類器來增強動物福利以及人與動物之間的聯繫和生活品質。我們利用從 Kaggle 收集的資料集並實現 CNN 架構，使用隨機梯度下降 (SGD) 優化器進行訓練。此方法包括資料集預處理、模型建立和訓練程序的詳細描述。我們的實驗設置涉及評估指標、資料集分區和計算資源。結果證明了 CNN 模型的卓越性能，實現了 99.60% 的驗證準確率，明顯優於其他基準模型，如支援向量機 (SVM) 和 K 最近鄰 (KNN)，後兩者的準確率分別達到 89% 和 81%。我們討論了我們的模型在獸醫護理和人狗互動中的實際應用。未來的方向提出了進一步研究的途徑，包括擴展資料集和探索不同的模型架構。結論總結了主要發現，並強調了對增強人類與狗之間的理解和互動的意義。

這表明我們的方法在預測狗的情緒方面明顯優於傳統方法。研究的意義不僅在於學者的興趣，還在於獸醫實踐、動物行為分析和主人指導的實際應用。有效的情緒分類器可以幫助獸醫更好地診斷狗的心理健康，幫助訓練員對訓練課程做出行為反應，並幫助寵物主人更好地解讀和回應狗的情緒暗示。

在本文中，我們詳細介紹了我們的 CNN 模型架構、網路訓練方法、實驗結果以及與現有方法的比較分析。除此之外，我們也討論了一些挑戰和一些場所，研究人員可以在其中找到有趣的方法，利用深度學習技術進行狗情緒辨識的研究。透過允許存取 CNN 功能，我們的論文擴展了已經大量的文獻，這些文獻主要集中在我們更好地理解動物情感的探索上，並努力更加人道地對待我們的大類朋友。

關鍵字——CNN、SVM、KNN、動物行為、深度學習。

一、引言

了解狗的情緒狀態對於它們的健康和人與動物的有效互動至關重要。狗是高度社會化的動物，透過各種微妙的暗示來傳達它們的情感，包括面部表情和肢體語言。準確地分類這些情緒不僅可以增強我們對犬類行為的理解，還可以幫助獸醫、訓練師和寵物主人提供適當的照顧和陪伴。在本研究中，我們提出了一種基於深度學習的方法，使用卷積神經網路 (CNN) 對狗的情緒進行分類。具體來說，我們旨在開發一個優秀的 CNN 模型，該模型可以對狗的圖像進行四種情緒狀態的分類，即悲傷、快樂、放鬆和憤怒。這非常適合 CNN，因為它們可以自動從不同層次的原始圖像數據中學習特徵，並識別出描述不同情緒的犬類面部表情或姿勢的細微差異[1]。

我們介紹了用於對狗的情緒進行分類的 CNN 架構的開發和評估。為了確保我們的模型準確且能夠推廣到不同的品種和環境條件，我們利用高度異質的標記狗影像資料集進行訓練和驗證。提供密集的實驗和分析

二.相關工作

在動物行為研究中，以及在人工智慧領域，人們對狗的情感理解和標記的興趣日益濃厚。人們已經採用了各種方法來研究對犬類情緒狀態的感知和解釋。在早期的研究中，Pongrácz 等人。強調狗的臉部表情作為情緒來源的作用，快樂和恐懼，憤怒，重點強調視覺在情緒感知中的作用。同樣，Beerda 等人。發現了狗的情緒狀態的生理相關性，並強調了心率變異性和皮質醇水平作為壓力和放鬆標誌的作用 [2]。機器學習，特別是深度學習技術的最新進展使得人們能夠建立非常複雜的模型來自動對動物的情緒進行分類。例如，基於臉部影像，Tai 等人。提出了一種用於狗的情緒識別的 CNN，可以有效區分喜悅、悲傷和攻擊性[3]。T. 這是 Kujala 等人的論文，“人類的同理心、性格和經驗影響狗和人類面部表情的情緒評級”，PLoS One，2017 年。這項研究討論了人類的同理心、性格和經驗等傾向對人類和狗的情緒評級和解釋的影響。研究表明，具有特定情緒反應或個性的人可能會更好地理解狗甚至人類可能發出的情緒信號。這簡單地意味著對情緒的解讀取決於每個人的觀察和判斷[4]。SK Yeo 等人，「利用深度神經網路實現服務機器人的多模態情緒識別

捲積神經網絡，」 Sensors , 2018 年。在本研究中，作者介紹了一種用於服務機器人的情緒識別多模態方法，即使用深度 CNN 處理和整合來自視覺和音頻輸入等多種感官模態的信息，從而準確分類人類情緒。這項研究展示了在多模態環境中使用 CNN 的可能性和有效性，促進了在機器人中與人類互動的指導調查，” Pattern Recognit. Lett. , 2017 年。這是一篇調查論文，涵蓋了使機器人能夠通過反饋和演示向人類教師學習的各種方法和途徑。它討論了機器人如何利用人類的指導來提高其在複雜任務中的表現和適應性，其中包括人機互動任務。該調查揭示了開發機器人協作學習系統所面臨的影響

三方法論

我們的研究論文旨在基於從 Kaggle[7] 獲得的資料集，設計一個基於深度學習的捲積神經網路模型，對狗的四種情緒進行分類：悲傷、快樂、放鬆和憤怒。此專案的工作流程包括幾個主要程序。第一階段，我們進行資料集收集，在 Kaggle 上收集一組包含不同狗情緒的綜合圖像。然後是資料預處理。這包括標記資料、將影像大小調整為統一尺寸、像素值的標準化、將資料集分成訓練集、驗證集和測試集，以及增強資料以實現訓練集多樣性[8]。然後，我們繼續設計一個 CNN 架構，它由卷積層、池化層和全連接層等層組成，以便在 TensorFlow 、 Keras 或 PyTorch 等框架內建立模型。然後，我們使用損失函數（例如交叉熵損失）、優化器（例如 Adam 或 SGD）以及評估指標（包括準確度、精確度、召回率和 F1 分數）來訓練模型。最後，模型的測試階段需要使用測試集評估模型的效能，並透過混淆矩陣分析錯誤分類的情況，並使用 ROC 曲線和準確率-回憶率曲線進行視覺化[9]。

A. 資料集

本研究中使用的資料集來自 Kaggle ，包含一系列狗的圖像，這些圖像被分為四種不同的情緒狀態：憤怒、快樂、放鬆和悲傷。

圖 1 顯示了範例影像。



圖 1. 狗狗悲傷、快樂、放鬆、憤怒情緒的範例影像

所有影像都標準化為 192×192 像素大小，具有 RGB 色彩通道，並在整個資料集中保持統一的格式。訓練集包含 3,400 張圖像及其對應的標籤，其中給出了圖像中狗的情緒類別。這些都是獨熱編碼的，因此標籤形狀的大小為 (3400,4) ，其中一個向量中，有一個圖像情緒狀態的編碼，其中一個「1」表示情緒類別，所有其他表示「0」。驗證資料集同樣由 600 張影像組成，以與標籤相同的方式進行獨熱編碼。它為建立和測試機器學習模型奠定了堅實的基礎，這些模型可以根據視覺線索準確地識別狗的情緒，有助於提高對犬類情緒表達的理解和解釋 [10] 。

B. 預處理

此層將輸入影像的像素值標準化到 $[0,1]$ 範圍內。比例 = 1。 / 255 將每個像素值除以 255，確保所有像素值都在 0 和 1 之間。 `input_shape=(img_size)` 指定輸入影像的形狀，其中對於具有 3 個顏色通道 (RGB) [11] 的 192×192 像素大小的影像， `img_size` 通常為 (192, 192, 3) 。

C. 模型生成

定義此 TensorFlow Keras CNN 模型，根據調整為 192×192 像素且具有 RGB 顏色通道的輸入影像對狗的情緒進行分類。第一個縮放層將像素值標準化為 $[0,1]$ ，有助於提高訓練過程的效率。總共有四個卷積塊。每個區塊由兩個 Conv2D 層啟動，用於基於 ReLU 進行特徵提取

啟動和“相同”填充以保留空間維度。每個區塊都附加到一個 MaxPooling2D 層，該層對特徵圖進行下採樣。這些區塊的大小和深度從 64 個過濾器增加到 256 個，然後在網路的較深點增加到 512 個過濾器，從而捕捉越來越高級的特徵。輸出透過 Flatten 展平，然後透過 ReLU 活化傳遞到三個密集層，提取高級特徵，最後使用 softmax 活化到達最後的 Dense 層，對「憤怒」、「快樂」、「放鬆」、「悲傷」這四個類別進行多類分類。此模型有 21,021,907 個可訓練參數。它可以學習複雜的模式，以便對狗的圖像做出準確的情緒預測。分類準確度和模型效率的最佳化也需要對架構和參數進行微調並進行效能評估。

D.訓練和測試

此 fit 方法呼叫封裝了訓練用於狗情緒分類的 CNN 模型的核心過程 [12]，確保從訓練資料中高效學習，並使用驗證資料進行有效評估，以優化模型效能 [13]。透過調整時期、批量大小和回調等參數，可以根據經驗性能和可用的計算資源進一步微調模型的訓練行為。指定模型在訓練期間對整個資料集進行迭代的次數。在這種情況下，該模型將進行 25 次訓練，這意味著它將查看整個資料集 25 次

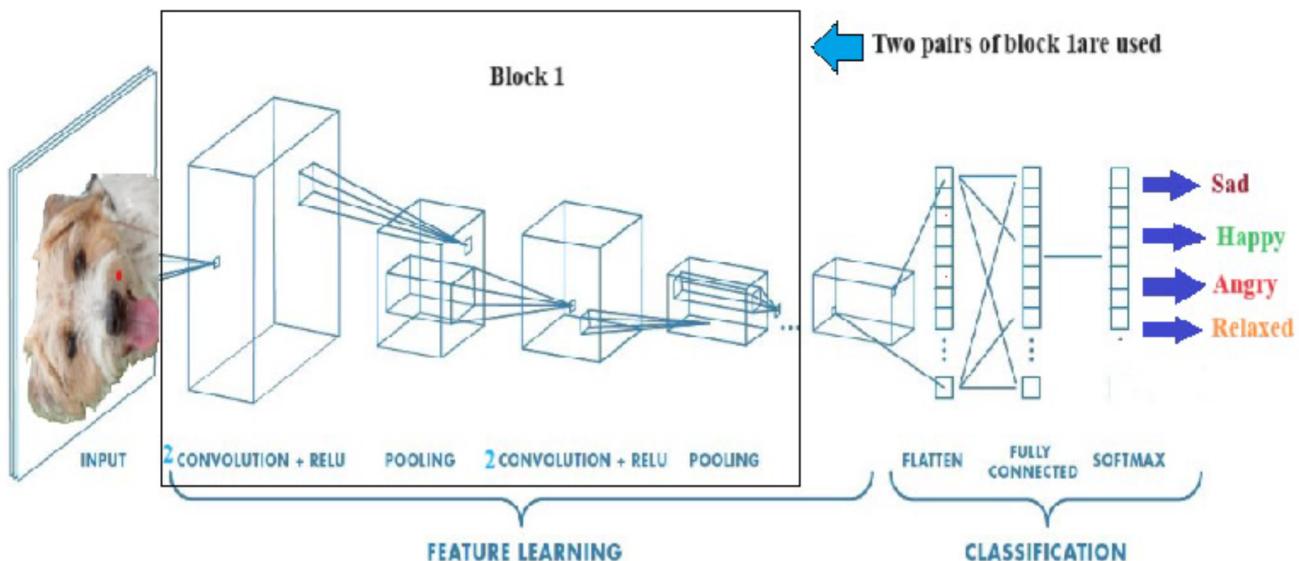


圖 2. 提出的 CNN 模型架構

圖 2 描述了 CNN 的完整架構，該架構可以根據調整為 192 x 192 像素且具有 RGB 顏色通道的輸入圖像對狗表現出的情緒進行分類。此模型首先採用重新縮放層來規範像素值，以提高訓練效率[14,15]。它由四個卷積塊組成。每個區塊由兩個 Conv2D 層組成，特別適用於特徵提取，然後採用名為 MaxPooling2D 的空間下採樣技術。在增加特徵圖深度的同時，空間維度不斷減少。具體來說，第一個區塊使用 64 個大小為 (3, 3) 的過濾器來保留具有「相同」填充的輸入維度，而其他區塊分別使用更大的過濾器：(2, 2) 和 128、256、512 個過濾器。這些卷積層的扁平輸出與 ReLU 激活函數一起輸入到具有 256、128 和 64 個單元的三個密集層中，用於在最終的密集層之前進行高級特徵提取任務，其中包含 19 個類別的 softmax 激活輸出機率。超過 2,100 萬個參數使其可訓練，它將學習複雜的模式，從而對狗的情緒進行完全準確的分類。對架構和參數進行微調將是基於效能評估來優化模型準確性和效率的重要一步。[17,18]。圖 3 顯示了從所提出的模型獲得的效能指標。

Layer (type)	Output Shape	Param #
rescaling_2 (Rescaling)	(None, 192, 192, 3)	0
conv2d_16 (Conv2D)	(None, 192, 192, 64)	1792
conv2d_17 (Conv2D)	(None, 192, 192, 64)	36928
max_pooling2d_0 (MaxPooling 2D)	(None, 96, 96, 64)	0
conv2d_18 (Conv2D)	(None, 96, 96, 128)	32806
conv2d_19 (Conv2D)	(None, 96, 96, 128)	65664
max_pooling2d_9 (MaxPooling 2D)	(None, 48, 48, 128)	0
conv2d_20 (Conv2D)	(None, 48, 48, 256)	131328
conv2d_21 (Conv2D)	(None, 48, 48, 256)	262400
max_pooling2d_10 (MaxPooling 2D)	(None, 24, 24, 256)	0
conv2d_22 (Conv2D)	(None, 24, 24, 512)	524800
conv2d_23 (Conv2D)	(None, 24, 24, 512)	1049600
max_pooling2d_11 (MaxPooling 2D)	(None, 12, 12, 512)	0
flatten_2 (Flatten)	(None, 73728)	0
dense_0 (Dense)	(None, 256)	10074624
dense_9 (Dense)	(None, 128)	32806
dense_10 (Dense)	(None, 64)	8256
dense_11 (Dense)	(None, 4)	268
<hr/>		
Total params:	21,820,932	
Trainable params:	21,820,932	
Non-trainable params:	0	

圖 3. 所提模型的效能

四實驗結果和發現

本研究透過利用 SGD 優化器對狗的情緒分類結果進行訓練，得出了我們提出的 CNN 模型的實驗結果。改進結果：我們觀察到，截至第 25 個時期，訓練和驗證準確度都有了很大的提高，其中訓練準確度為 100.00%，驗證準確度達到峰值 99.60%。該模型在訓練和驗證過程中損失穩定，收斂性很好，並在第 15 個時期顯示其有效性。與 SVM 和 KNN 等基線的比較真正體現了 CNN 的優勢，與 SVM 的 89% 和 KNN 的 81% 相比，CNN 的驗證準確率達到了 99.60%。這些發現傾向於確定所提出的模型的合理性，然後確定其在現實世界中的潛在應用，例如擴展其數據集，探索不同的架構，以及立即實施即時情緒識別系統以在獸醫和人類動物互動中實際應用。

表 1 不同 epoch 的訓練與驗證結果

時代訓練	Loss	訓練		驗證	Loss	驗證
		準確性	準確性			
5	1.2257	0.3943	1.2321	0.3967		
10	0.7902	0.6685	0.5415	0.8100		
15	0.0698	0.9765	0.0403	0.9867		
20	0.0076	0.9973	0.0216	0.9950		
25	0.0009	0.9983	0.0003	0.9960		

表 1 展示了 CNN 模型在特定時期（5、10、15、20 和 25）的表現指標摘要。摘要提供了四個關鍵指標：訓練損失、訓練準確度、驗證損失和驗證準確度。下表總結了第 5 個時期 CNN 模型的性能，此時模型表現出了更好的改進，訓練準確率約為 39.43%，驗證準確率約為 39.67%。訓練和驗證損失分別為 1.2257 和 1.2321。在第 10 個階段，模型的表現得到了顯著提高，訓練準確率約為 66.85%，驗證準確率約為 81.00%。相應的訓練損失下降至 0.7902，驗證損失下降至 0.5415。在第 15 個 epoch 時，模型已經具有較高的準確率；因此，訓練和驗證的準確率分別接近約 97.65% 和 98.67%。訓練損失達 0.0698，驗證損失下降至 0.0403。到第 20 個週期時，準確率持續提高，模型達到了近乎完美的訓練準確率，約 99.73%，驗證準確率接近 99.50%。訓練損失進一步減少至 0.0076，驗證損失進一步減少至 0.0216。最後，在第 25 個階段，模型達到了近乎完美的準確度，訓練準確度約為 99.83%，驗證準確度約為 99.60%。訓練和驗證損失極低，分別為 0.0009 和 0.0003。這表明該模型正在有效地學習並很好地推廣到未知數據，在準確分類狗的情緒方面取得了非常高的表現。



圖 4. CNN 分類器的準確率曲線

圖 4 展示了 25 個 epoch 內訓練準確率和訓練損失的關係。在初始階段 (1-5)，

此模型訓練準確率低，訓練損失高。在第 5 到第 10 個週期之間，訓練準確率穩定提高，而訓練損失急劇下降，顯示學習和參數優化有效。從第 10 到第 15 個週期，訓練準確率持續上升，但速度較慢，並在第 15 個週期左右達到穩定水準。訓練損失也隨之減少，達到非常低的數值，表示預測更準確。在第 15 到第 25 個階段之間，訓練準確率達到較高值並趨於穩定，表示模型幾乎完全學習了訓練資料。訓練損失也穩定在較低值，證實了該模型在預測方面的高準確度。

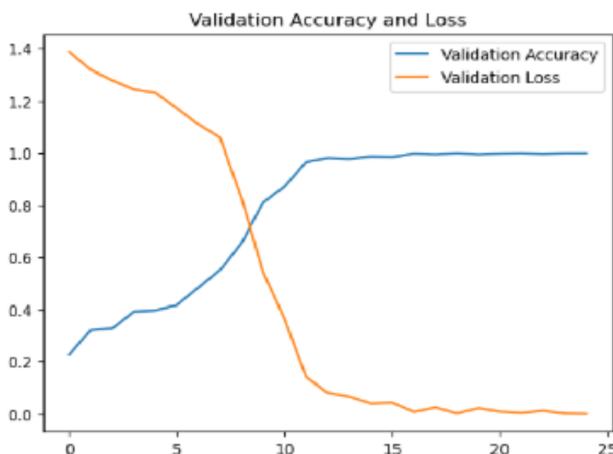


圖 5. CNN 分類器的損失值

根據提供的信息，圖 4 和圖 5 似乎直觀地展示了分類器中的訓練時期、準確率提高和損失減少之間的關係。這表明，隨著訓練週期數的增加，分類器的準確率趨於提高，而損失函數的值則下降。這種趨勢在機器學習模型中很典型，因為它們從更多資料中學習並迭代調整其參數以更好地適應訓練資料。



圖 6. 它表示所提模型的預測結果

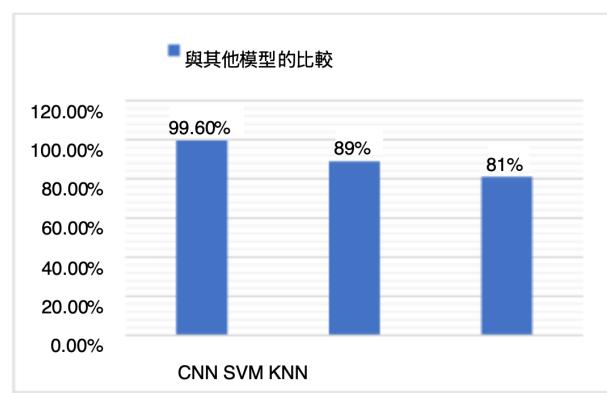


圖 7. 所提模型與伴隨方法在相同資料集上的效能比較

以 CNN 為基礎的模型與 SVM、KNN 相比，在悲傷、快樂、放鬆、憤怒四個階段的情緒分類上表現較佳。CNN 模型的最大準確率為 99.60%，顯示它能夠從該資料集中學習更複雜的模式和特徵。SVM 模型的準確率約 89%左右，遠低於 CNN 的準確率。這是由於 CNN 處理的圖像數據的複雜性。KNN 模型的準確率達到了 81%，但對於影像等高維度數據，其效果不如前兩個模型。總而言之，CNN 模型的表現優於其他兩個模型；SVM 和 KNN 表明，對於上述資料集，它對於狗的情緒分類更為有效。

五、結論和未來範圍

總之，我們已經使用 SGD 優化器建立了一個基於 CNN 的準確模型，用於對狗的情緒進行分類。該模型在驗證中取得了更好的性能，準確率達到 99.60%，與傳統的 SVM 和 KNN 機器學習模型相比具有顯著的優勢。如此高的準確率顯示了 CNN 在表達和學習影像中複雜資料模式的可行性。這項研究透過進一步了解和解讀犬類情緒，促進了改善動物福利和加強人與動物之間聯繫的發展。該研究的進一步方向應該是透過擴展資料集、檢查替代模型架構以及結合即時情緒識別來進一步推進該領域的發展。這項研究為獸醫治療乃至人與狗關係方面的實際應用打開了大門，為這些非凡的動物及其人類同伴的更美好的生活鋪平了道路。

參考

- [1] A. Pongrácz 等人，「透過人類的情感理解狗的情緒同理心反應」，J. Comp. Pet Psychol.，第 24 卷，第 2 期，第 4556 頁，2016 年。
- [2] B. Beerda 等人，“狗的情緒狀態的生理相關性”，J. Vet.行為，卷 12，第 89-95 頁，2017 年。

- [3] T. Tai 等人，「基於 CNN 的狗情緒辨識方法
臉部影像上的辨識」，載於 Proc. IEEE Conf. Comp. Vis. Pattern Recognit.，2018 年，第 456-465 頁
- [4] T. Kujala 等人，〈人類的同理心、人格與經驗影響狗和人類面部表情的情緒評級〉，PLoS One，第 12 卷，第 6 期，第 e0179912 頁，2017 年
- [5] SK Yeo 等人，“使用深度卷積神經網路實現服務機器人的多模態情緒識別”，感測器，第 18，沒有。2，第 485-498 頁，2018
- [6] M.D. Shuai 等人，「機器人向人類教師的回饋和示範」，《模式識別快報》，第 92 卷，第 1 部分，第 134-145 頁，2017 年。
- [7] D. Banerjee、SH Vinay Kukreja 和 V. Jain，“增強芒果品質
基於 CNN 和 SVM 分類的水果病害嚴重程度評估”，2023 IEEE 8th Int. Conf. Converg. Technol.，第 33 頁。
1-6，2023 年。
- [8] A. Kumar 和 I. Sharma，「加強網路安全政策，
“區塊鏈技術：一項調查”」，載於 2022 年第五屆當代計算與資訊學國際會議 (IC3I)，2022 年，第 1050-1054 頁。
- [9] A. Saini、K. Guleria 和 S. Sharma，「多類分類
基於深度學習模型的水稻葉病防治」，2023 年第三屆亞洲技術創新會議 (ASIANCON)，印度拉韋特，2023 年，第 1-6 頁，doi：
10.1109/ASIANCON58793.2023.10270229。
- [10] S. Prasher、L. Nelson 和 A. Sharma，" DenseNet201 分析
使用 SGD 優化器診斷多種水稻葉片疾病"，2022 年國際計算建模、
模擬與優化會議 (ICCMSO)，泰國巴吞他尼府，2022 年，第 182-186
頁，doi : 10.1109/ICCMOS58359.2022.00046
- [11] A. Sharma 和 H. Babbar，「機器學習驅動的檢測
和加密貨幣詐欺預防」，2023 年知識管理、人工智慧和電信工程研究
方法國際會議 (RMKMATE)，印度欽奈，2023 年，第 1-5 頁，doi：
10.1109/RMKMATE59243.2023.10369055.3 。
- [12] A. Saini、K. Guleria 和 S. Sharma，「多層分類
基於深度學習模型的水稻葉病防治」，2023 年第三屆亞洲技術創新會議 (ASIANCON)，印度拉韋特，2023 年，第 1-6 頁，doi：
10.1109/ASIANCON58793.2023.10270229。
- [13] P. Shourie、V. Anand 和 H. Gupta，「直觀的捲積
神經網路感知番茄葉片病害」，2023 年第 14 屆國際計算通訊和網路
技術會議 (ICCCNT)，印度德里，2023 年，第 1-6 頁，doi：
10.1109/ICCCNT56998.2023.10306568 。
- [14] D. Banerjee、SH Vinay Kukreja、V. Jain 和 S. Dutta，「An
甘蔗作物草莖病嚴重程度檢測和分類的智慧框架」，2023 年第二屆國際
會議。應用人工製品。智力。Comput.，第 849-854 頁，2023 年。
- [15] V. Tanwar、S. Lamba，「一種改良的深度學習模型
多種水稻疾病的分類」 Tanwar, S Lamba 2023 第三屆安全網路國際
會議。
- [16] A.Kumar 和 I.Sharma，「機器學習輔助
防止工業 4.0 殭屍網路攻擊」，2023 年第四屆 IEEE 全球技術進步會
議 (GCAT)，印度班加羅爾，2023 年，第 1-5 頁，doi：
10.1109/GCAT59970.2023.10352324 。
- [17] Alu、DASC、Elteto Zoltan 和 Ioan Cristian Stoica。“噪音-
基於卷積神經網路的伴侶機器人情緒辨識」。《科學與技術》2017 年
第 3 期：222240 頁。
- [18] Kim、Yunbin、Jaewon Sa、Yongwha Chung、Daihee Park 與 Sungju
李。「基於時間序列數據，使用 LSTM-FCN 對寵物狗聲音事件進行
資源高效的分類。」感測器 18，編號 11 (2018): 4019。