

聲紋辨識用於輔助母豬分娩照護系統之研究⁽¹⁾

賴佑宜⁽²⁾⁽⁶⁾ 王斌永 林正斌⁽³⁾ 李秀蘭⁽⁴⁾ 沈煜棠⁽⁵⁾

收件日期：111 年 11 月 16 日；接受日期：112 年 9 月 25 日

摘 要

為避免母豬分娩後壓死剛出生的仔豬，本研究利用聲紋及影像之收集及分析，期待能開發出讓仔豬有時間避開將躺下的母豬且提醒母豬之智慧型母豬分娩照護系統。本試驗於母豬分娩舍建置母豬分娩收音監測設備，並依母豬鳴聲聲紋特性，開發出豬鳴聲訊號擷取軟體，收集即將分娩母豬、剛出生仔豬隻叫聲及仔豬被壓前的叫聲，並透過網際網路將聲紋資料傳輸至儲存之資料庫中進行分析及運用。所收集的聲紋經偵測後採梅爾倒頻係數 (Mel-Frequency Cepstral Coefficient, MFCC) 進行特徵向量比對及辨識演算，去除非上述擬收集的三種聲紋。結果顯示：經 MFCC 系統篩選出之 8,000 筆豬聲紋資料中，顯示母豬分娩前之聲頻均集中在 4,000 Hz 以下，剛出生仔豬叫聲聲頻則為 4,000 – 8,000 Hz，仔豬被壓時叫聲的聲頻則介於 2,000 – 4,000 Hz 間。綜合上述，母豬分娩聲頻及仔豬被壓之聲頻仍有重疊。因此，未來可利用影像之辨識區分二者聲頻及母豬將壓到仔豬之行為，進而開發出母豬分娩仔豬防壓之照護系統。

關鍵詞：豬場、機器人、人工智慧、分娩。

緒 言

2010 年全球養豬頭數約 9.7 億頭，至 2015 年達 9.9 億頭後，均維持在 9.7 億頭左右的在養頭數，惟 2019 年爆發非洲豬瘟疫情，全球養豬頭數頓時降至 8.5 億頭，大幅減少約 12%，截至 2021 年，全球約有 7.5 億頭豬 (Statista, 2022)，豬肉亦是消費量第二大的肉類，隨著人口增長，豬肉需求將持續增加，為滿足日益增長的肉品需求，及解決逐漸增加的人力成本，近年養豬產業積極發展精準畜牧飼養 (Precision livestock farming, PLF)，透過非侵入式監控設備、行為偵測及數據分析應用，發展即時監測或自動化設備，以增加生產效率、提升動物健康以協助農民管理牧場。

農業朝向集約化發展，導致對於經濟動物質量控制和重視動物福利的需求增加，可追溯性的動物識別日趨重要，個別自動識別是落實個體化管理在疾病檢測、控制和提升產品質量的重要步驟 (Marsot *et al.*, 2020)。在改善動物福利和提高養殖效率方面，利用物聯網 (Internet of Things, IoT)、邊緣計算 (Edge Computing)、雲端計算 (Cloud Computing) 和數據驅動 (Data Driven) 等技術，已經可以將物聯網傳感器所獲得的畜禽生物和環境信息串連到管理系統 (Iwasaki *et al.*, 2019)。所以養豬生產系統自動化與智能化為全球養豬 30 大國 (臺灣位居第 19 名) 必然趨勢，如同丹麥、德國與荷蘭等國應用自動感知、辨識、偵測、監控、自動控制以及數據推演等智慧化系統，建構智能化豬舍，提升養豬生產效率。

我國農業特性與勞動力環境，包括農業環境條件不佳、工作時段不固定、農工薪資所得低、農事操作具有專業技術、經營規模小及勞動環境倚賴農工等因素，以致青壯年從農意願低、勞動力老化，農業缺工問題叢生 (曾及張, 2019)。養豬產業面臨國內養豬場人力老化的問題 (從業人員平均年齡超過 62 歲)，在智能化豬舍環境感測、精準管理及決策系統建置下，針對生長豬的精準飼養管理、種豬發情配種分娩哺乳、保育豬生長追蹤系統以及人工智慧化的養豬專家系統等方向，可以解決產業缺工與人力老化等問題。

根據 Research and Markets (2020) 報告指出，2020 年全球畜牧養殖智慧化市場規模約 31 億美元，預計以年複

(1) 農業部畜產試驗所研究報告第 2765 號。

(2) 農業部畜產試驗所北區分所。

(3) 農業部畜產試驗所飼料作物組。

(4) 農業部畜產試驗所技術服務組

(5) 財團法人工業技術研究院智慧微系統科技中心。

(6) 通訊作者，E-mail: laiyou@mail.tlri.gov.tw

合成長率 (Compound annual growth rate, CAGR) 9.0% 成長, 2025 年市場規模將達到 48 億美元。主要促進因素包含對於物聯網、智慧化監測設備、人力替代及自動化之需求增加, 以及日益重視即時監測和早期疾病檢測, 對蛋白質需求不斷增長等。至於 COVID-19 疫情對於全球畜牧養殖智慧化市場的影響主要在 2020 年第 1 至第 3 季, 主因係各領域的業主缺乏添購新設備之意願, 然而此現象在 2020 年底開始復甦, 並於 2021 年下半年逐步回復常態。

畜牧產業在全球經濟產業鏈當中扮演相當重要的角色, 由傳統的自然放牧與圈養, 直到現代因科學技術進步, 在生物學、遺傳學以及工業機械的蓬勃發展之下, 家畜的養殖、繁衍、數量管理等均有突破性的發展, 但在家畜之健康管理當中, 往往須透過人力來做初步的監控, 而根據研究指出, 家畜之鳴叫聲也可反映其生理狀況 (Jahns *et al.*, 1998)。

Shen *et al.* (2021) 之研究報告指出, 豬咳嗽被認為是呼吸系統疾病最常見的臨床癥狀, 因此通過監測和建立豬呼吸系統疾病的早期預警系統識別其咳嗽音是很重要。該研究提出了一種新的融合特徵, 即梅爾頻率倒譜係數卷積神經網路 (Mel-frequency cepstral coefficients-convolutional neural network, MFCC-CNN), 用以提高豬咳嗽的識別準確率, 研究人員分別使用歸一化指數函數 (softmax) 和線性 (linear) 支援向量機 (SVM) 分類器進行分類。結果顯示, 使用 MFCC-CNN 特徵的分類器性能明顯優於使用 MFCC 特徵的分類器。Softmax 和 SVM 分類器的 F1 得分 (機器學習模型精確度指標分數) 分別提高了 10.37% 和 5.21%, 咳嗽準確率分別提高了 7.21% 和 3.86%。

Wu *et al.* (2022) 認為豬的聲音可以反映饑餓、疼痛或情緒等各種狀態, 是重要標誌之一, 可直接顯示豬隻生長和健康情況。現有的聲音辨識方法通常從頻譜特徵 (Spectrum Signature) 開始, 使用頻譜圖 (Frequency Spectrum) 來對應不同的分類, 雖然效果很好, 但或許並非解決具有一維特徵輸入的此類研究的最佳方法。為了更準確掌握豬的情況, 及時採取措施保證豬的健康情況, Wu *et al.* (2022) 提出了一種基於信號頻譜和聲紋 (Voiceprint) 雙重作用的豬聲音分類方法。頻譜圖可以可視化有關不同時間段下聲音特徵的資訊, 引入音訊數據, 將模型輸入的頻譜圖特徵和音訊時域特徵相互補充, 傳遞至預先設計的併行網路結構中, 再選擇結果最好的網路模型和分類器進行組合。結果豬聲音分類的準確率達到 93.39%, 曲面下面積 (Area under curve, AUC) 也達到 0.99163, 證明該方法的優越性。

Shi (2022) 發表一種改進的豬異常聲音辨識機器學習模型, 作者同樣支持豬的不同聲音信號反映當前生理健康和情緒狀態的觀點, 因此亦以豬咳嗽作為識別標的, 首先, 分析豬叫信號的時域 (time domain) 和頻域 (frequency domain), 搭建了豬音訊採集的硬體系統, 並針對收集到的豬音訊信息進行降噪。活豬聲音端點檢測採用雙閾值終點檢測法 (double threshold endpoint detection method), 然後利用梅爾頻率倒譜係數提取活豬音訊信號的特徵。在豬咳嗽的識別中, 利用隱藏式馬可夫模型 (hidden Markov model) 通過機器學習提高識別精度和效率。實驗顯示, 本研究的識別方法確實能夠準確識別豬咳嗽音。

以牛隻畜養為例, 在牛隻身體異常或是分娩期間, 其鳴聲將產生變化, 因此畜牧業者透過牛隻鳴聲之變化將可以作為健康管理初步監控指標。隨著近年來 AI 人工智慧的興起, 影像辨識、聲紋辨識等演算法迅速發展, 且根據行政院農業委員會 (2016) 推動之智慧農業 4.0 計畫, 將從人、資源以及產業三方面進行優化, 實現「建置農業生產力知識及服務支援體系」之標的。

本研究透過聲紋大數據資料分析, 針對仔豬受壓迫的尖叫聲監測, 若發現異常狀況, 系統主動通報, 以利飼養人員針對突發事件進行處理, 減少損失。

材料與方法

I. 研究場域

本研究針對母豬分娩舍建置智慧聲音監控系統, 場域為農業部畜產試驗所 (以下簡稱畜試所) 畜產經營組二股母豬分娩舍。

II. 研究方法

首先利用母豬分娩舍建置收音監測設備, 進行聲音收集及監控, 可以透過網際網路將聲紋資料傳輸至畜試所指定之資料庫進行儲存。在母豬分娩舍所收集的仔豬叫聲聲紋, 經過建構預測分析模型後, 套用於現場即時監控, 當監測到仔豬被壓的情況, 立即驅動分娩舍智慧型機械設施調整母豬姿勢或位置, 使其不致壓斃仔豬, 降低其窒息的風險, 資訊系統規劃架構如圖 1 所示。本研究於系統建置過程中搭配影像系統, 但目前僅作為人工辨識豬隻行為作為聲紋研究使用, 並錄製相關影像資訊資料作為後續影像分析研究使用。

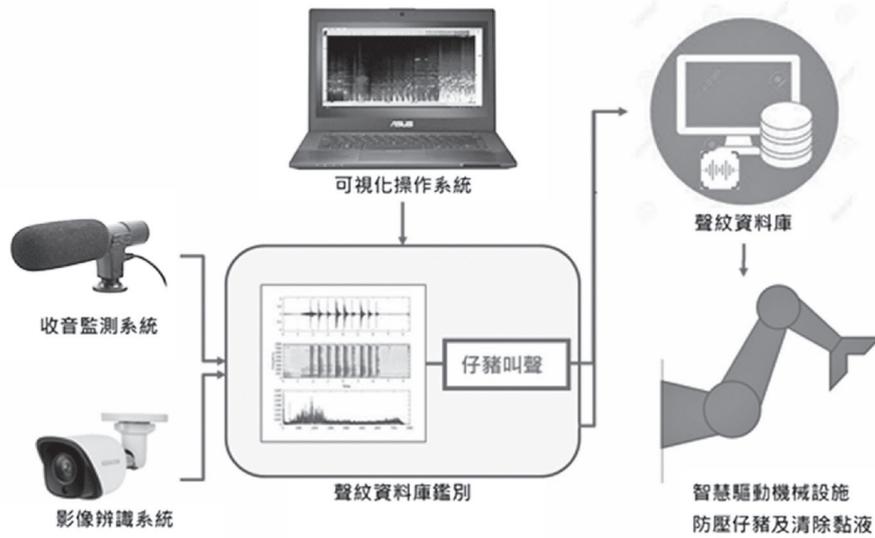


圖 1. 母豬分娩舍智慧型聲音監測系統架構示意圖。

Fig. 1. Schematic diagram of the intelligent monitoring system of the sow farrowing house.



圖 2. 母豬分娩舍聲音監測建置圖。

Fig. 2. Construction diagram of sound monitoring in the sow farrowing house.

母豬分娩舍所建置之的仔豬叫聲收集聲紋裝置設備如圖 2 所示，麥克風主要安裝於母豬分娩欄上方（圖 2 圈內），約 15 – 20cm 的地方，方便收集母豬和仔豬的叫聲，而測試的分娩欄位上方均設置一隻麥克風，總共架設 4 隻麥克風進行聲紋相關收集。

豬隻聲紋主要研究步驟為先作端點偵測（Entropy Energy），利用梅爾頻率倒譜係數（Mel-frequency cepstral coefficients, MFCC），最後再進行特徵向量比對，進行辨識。辨識流程如圖 3 所示。

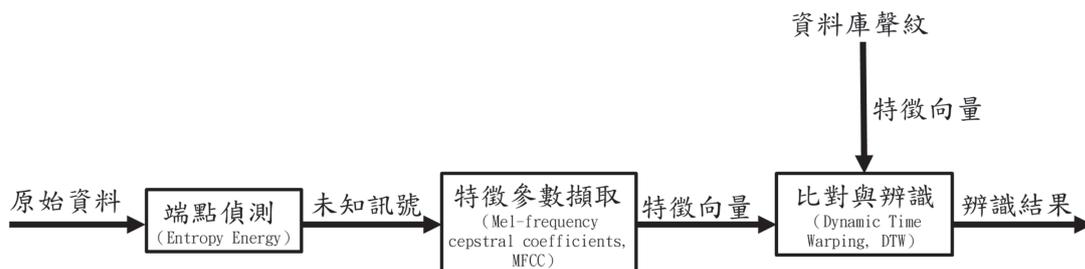


圖 3. 聲紋辨識流程。

Fig. 3. Voiceprint recognition process.

母豬分娩舍之錄音系統自動儲存錄音資料為 (Waveform Audio File Format, wav) 檔，為有效進行自動聲紋資料擷取，首先利用聲音頻譜的能量強度的熵 (entropy) 將有事件的聲音時間段挑出來，entropy 公式如下：

$$H = - \sum_{k=1}^N p_k \log(p_k)$$

H 為聲音頻譜的資訊熵 (information entropy)，也就是能量強度的熵，而 p 為該聲音事件發生的機率，k 為聲音事件種類的數量。當聲音在頻譜上出現事件時 (聲音能量在頻譜上較集中)，所計算出來的熵會較小 (如圖 4 所示)，故可以設定門檻，將低於門檻的聲音挑出來，當作聲音的端點。現階段軟體設計可自動於音訊檔儲存資料夾中自動批次載入檔案並進行聲音事件偵測，偵測得之資料則進行後續頻譜特性確認處理。

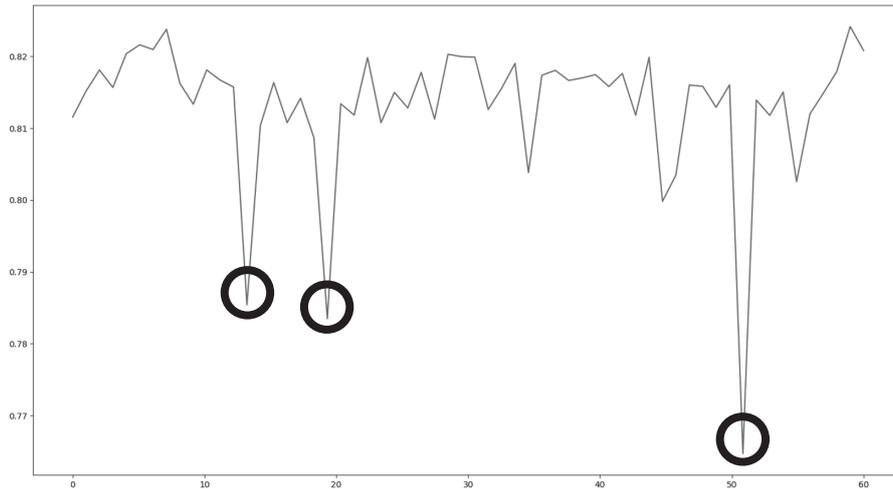


圖 4. 一分鐘內每秒的譜熵 (spectral entropy) 數值範例。

Fig. 4. Examples of spectral entropy values per second for one minute.

在母豬分娩舍之聲音事件，除了背景噪音外主要有母豬、仔豬、人聲、以及敲擊 (撞擊) 脈衝聲，因此針對聲音事件偵測之擷取輸出進行進一步分析，將人聲與敲擊 (撞擊) 脈衝聲排除，最後留下母豬與仔豬之聲訊並加以輸出該段聲訊與時頻譜圖。經過此程序擷取出之資料將用於建置豬鳴聲資料庫，以便進行機器學習及進一步分辨豬鳴聲情境。

本研究所選用的梅爾倒頻譜係數是一種常用的音訊特徵係數，梅爾倒頻譜係數被廣泛的運用在音訊及音樂資料庫的分類決策，除此之外，MFCC 也常應用於語音辨識以及生物聲音的辨識上，而其優點就是 MFCC 的辨識率相當的高，從蔣 (2008) 的研究得知，在語者辨識技術上，採用梅爾倒頻譜參數的方式有最好的辨識結果，即使是處於有噪音干擾導至有雜訊的音訊中，也能維持較佳的辨識率，因 MFCC 採用高頻少取低頻多取的方法，所以大多處於高頻的雜訊，也會因此被過濾掉。因本研究豬舍環境較為吵雜，為有效取得頻譜資料，建置聲紋資料庫，進而分析演算。

結果與討論

母豬分娩舍錄音資料存放電腦為主要平台，依豬鳴聲聲紋特性開發豬鳴聲訊號擷取軟體，並於該平台上直接運作進行豬鳴聲訊號擷取，目前已取得超過 8,000 筆豬鳴聲資料，當中以仔豬之豬鳴聲居多，相關豬鳴聲判別如圖 5。黃 (2011) 進行語音辨識相關研究，開發一套適合多種夜行性動物的辨識系統，將各個聲音樣本濾去雜訊、修正高頻，經特徵音訊切割，特徵萃取等程序，作為系統核心辨識之用，進行決策樹與模糊格推論之分類模組建置。本研究利用頻譜係數分析法，建立不同類型豬鳴叫頻譜背景資料，提供後續智慧化判定。

本研究參考 Wu *et al.* (2022) 研究，聲音樣本數據集記錄了四種分類，進食、正常、受驚和焦慮，採樣頻率為 48,000 Hz，及分類出 MFCC 基本特徵。母豬和仔豬同欄，所以收音的過程中會同時收到仔豬和母豬的聲音，因此需要將母豬和小豬的聲音進行分離才可以得到正確的仔豬聲音。於是利用錄影畫面來判斷各種狀態下的母豬和仔豬的聲音，如圖 6。圖 6 (a) 和圖 6 (b) 均為母豬的叫聲，而圖 6 (c) 和圖 6 (d) 為仔豬的聲音，如圖 6 所示，即時在生產前，母豬的叫聲頻率範圍均集中在 4,000 Hz 以下，而仔豬的聲音在 4,000 - 8,000 Hz 會有明顯的聲紋能量，因此，透過頻率響應高低頻的能量密度，可以正確區分出母豬和仔豬的聲音。

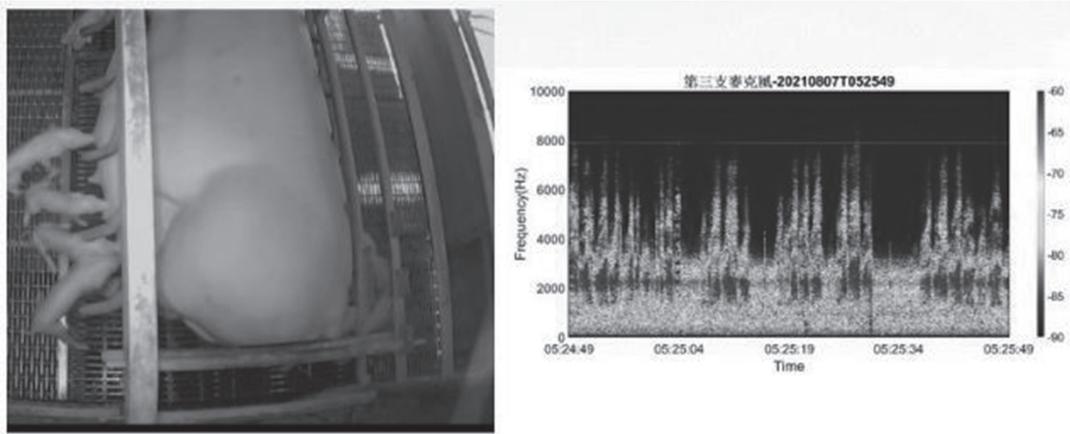


圖 5. 仔豬出生後哺乳之錄影畫面（左）及鳴叫聲之時頻譜圖（右）對照。

Fig.5. Comparison of the video image (left) and the frequency spectrum (right) of the piglet foraging after birth.

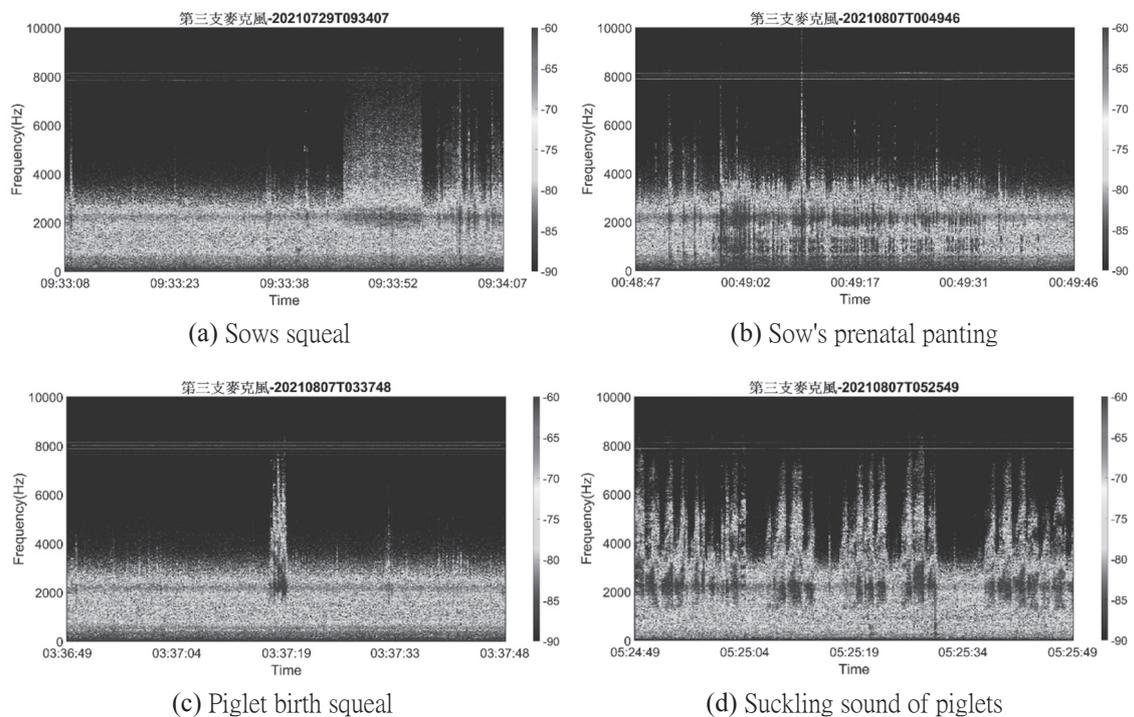


圖 6. 不同情境下母豬和仔豬的聲音。

Fig.6. Sow and piglet sounds in different situations.

在日常養豬過程中，仔豬剛出生的時候，會有少數小豬被母豬壓死，尤其是小豬剛出生 1 – 5 天，小豬行動還不靈活的時候。而有時候飼養不當還會咬死小豬，造成經濟損失。母豬壓死小豬的這種情況一般都是在母豬站起來喝水、排便、排尿後躺下時，或者是母豬翻身時造成的。寒冷的季節母豬壓死小豬的機率也會比往常高許多，這是因為天氣寒冷所以小豬特別喜歡靠在母豬身邊睡覺取暖。

一般仔豬聲紋和仔豬被壓聲紋的時頻譜圖如圖 7 所示，仔豬聲紋多為短促的叫聲，在時頻譜圖尚可看出較為明顯的時間差距，而仔豬被壓時會發出急促且連續的聲音，所以可以判斷為仔豬被壓。在母豬分娩舍所收集的仔豬叫聲聲紋，經過建構預測分析模型後，套用於現場即時監控，當監測到仔豬被壓的情況（圖 7）。當無法以聲音或影像辨識方式偵測到壓到仔豬，而遠端人員透過監控系統覺得需要的時候進行遙控操作，被壓仔豬只有極短黃金時間的拯救機會。

未來將加入影像分析之功能，透過影像與聲紋之收集、辨識，全天候輔助監控、觀察、偵測母豬接近分娩前之臨床徵兆，應用於母豬分娩舍中，即時通知管理人員提供適當協助；於分娩後哺乳階段，透過智慧型作動輕拍裝置（圖 8）將可有效避免壓死仔豬之情況，藉以降低損失、提升仔豬育成率。目前可暫時判定分類之聲紋資料有仔豬聲音、母豬聲音、母豬分娩前喘息聲、人聲、器械敲擊聲音和小豬覓食聲音，而仔豬被壓的聲音和母豬聲

音太過近似，同樣落在 2,000 – 4,000 Hz，相似頻譜，尚無法被有效區分，未來將採影像辨識輔助人工智慧判斷，Jahns *et al.* (1998) 指出可利用聲頻的差異可辨識動物的狀況，Marsot *et al.* (2020) 也指出豬可利用卷積神經網絡 (convolution neural networks) 等辨識系統辨別豬體等外觀之差異。因此，未來透過 AI 大數據分析，透過聲紋及影像辨識雙向管道，增加人工智慧的精準度，母豬及仔豬可得到較好的照護，除促進動物福祉外，亦能增加離乳仔豬存活率。

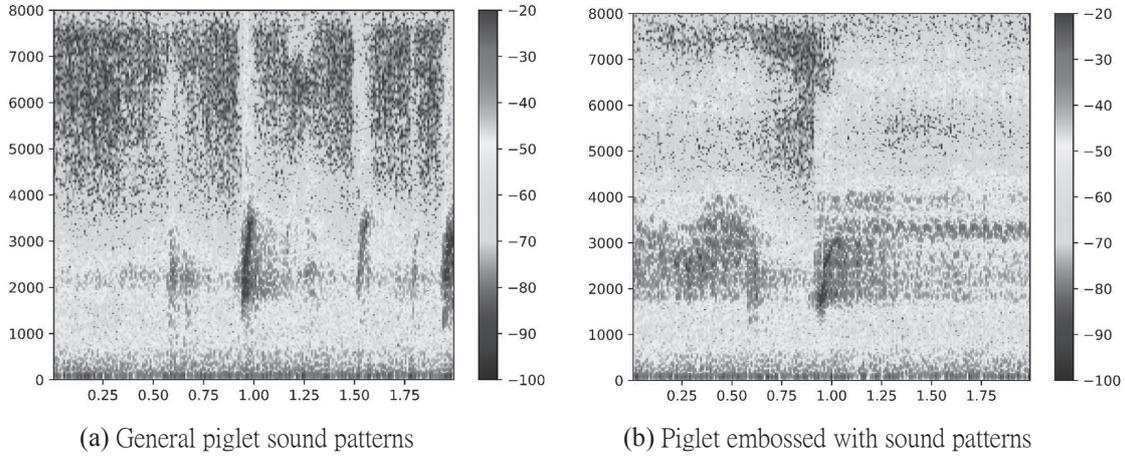


圖 7. 仔豬鳴叫聲頻譜圖。

Fig.7. Spectrogram of piglet chirping sound data.



圖 8. 智慧型作動輕拍裝置圖。

Fig.8. Smart motion tapping device diagram.

結 論

本研究針對仔豬不同狀況鳴叫聲提出了基於使用 AI 搭配音訊特徵 (Mel-scale Frequency Cepstral Coefficients, MFCC) 以用作分類的方法，不同於當前的影像辨識分類方法，礙於本系統目前所開發的技術，仔豬被壓的聲音和母豬聲音頻譜太過近似，同樣落在 2,000 – 4,000 Hz 的頻譜，另外目前仔豬被壓的樣本數太少 (目前僅有一有效樣本)，透過訓練學習無法判定為有效，尚無法精準辨識，未來研究必須投入影像研究輔助辨識，因聲音會受雜訊影響，而影

像會有拍攝死角，兩者相輔相成，除強化智慧型聲音監控系統準確度外，未來亦可將兩者整合成物聯網系統之一環，可協同作業，增加管理效率。

參考文獻

- 行政院農業委員會。2016。以智慧科技邁向台灣農業 4.0 時代。
<https://www.coa.gov.tw/ws.php?id=2505139>
- 曾康綺、張惠真。2019。農業季節性缺工 2.0 措施之農業人力調度個案研究－以臺中農業技術團為例。臺中區農業改良場研究彙報 144：35-49。
- 黃振榮。2011。夜行性動物聲音收錄與辨識系統。行政院農業委員會林務局保育研究系列 99-21 號。
- 蔣明達。2008。以時間領域濾波方式來計算梅爾倒頻譜參數的語者識別。中華技術學院電子工程研究所碩士班碩士論文，台北市。取自 <https://hdl.handle.net/11296/85v879>。
- Iwasaki, W., N. Morita, and M. P. B. Nagata. 2019. Iot sensors for smart livestock management. Elsevier, pp. 207-221.
- Jahns, G., W. Kowalczyk, and K. Walter. 1998. Sound analysis to recognize individuals and animal conditions. XIII CIGR Congress on Agricultural Engineering, 2-6 February 1998.
- Marsot, M., J. Mei, X. Shan, L. Ye, P. Feng, X. Yan, C. Li, and Y. Zhao. 2020. An adaptive pig face recognition approach using Convolutional Neural Networks. *Comput. Electron. Agric.* 173: 105386.
- Research and Markets. 2020. Global Precision Livestock Farming Market 2020-2024. <https://www.researchandmarkets.com/reports/5178373/global-precision-livestock-farming-market-2020>.
- Shen, W., D. Tu, Y. Yin, and J. Bao. 2021. A new fusion feature based on convolutional neural network for pig cough recognition in field situations. *Inf. Process. Agric.* (8): 573-580.
- Shi, Y. M. 2022. An improved machine learning model for pig abnormal voice recognition. *J. Computers* 33(6): 155-166.
- Statista. 2022. Number of pigs worldwide from 2012 to 2021 (in million head) . <https://www.statista.com/statistics/263963/number-of-pigs-worldwide-since-1990/>.
- Wu, X., S. Zhou, M. Chen, Y. Zhao, Y. Wang, X. Zhao, D. Li, and H. Pu. 2022. Combined spectral and speech features for pig speech recognition. *PLoS ONE* 17(12): e0276778. <https://doi.org/10.1371/journal.pone.0276778>.

Research on voiceprint recognition for assisted sow parturition care system ⁽¹⁾

Yui-I Lai ⁽²⁾⁽⁶⁾ Bin-Yong Wang Jeng-Bin Lin ⁽³⁾ Hsiu-Lan Lee ⁽⁴⁾ and Yu-Tang. Sheen ⁽⁵⁾

Received: Nov. 16, 2022; Accepted: Sep. 25, 2023

Abstract

The purpose of this project will be to develop a smart farrowing care system for piglets by collecting the voiceprints of sow farrow, piglet and piglets' cry. The voiceprint data would be transmitted to the database designated by the Mel-frequency cepstral coefficient (MFCC) through the Internet for storage. Those voiceprints of this project were collected in the sow farrowing house and were used for on-site real-time monitoring after constructing a predictive analysis model. The voiceprints of sow farrow, piglet and piglet's cry will be transmitted and tested by MFCC when those voiceprints were collected. The data showed those of the sow farrow and piglets' cry voice points below 4,000 Hz, 4,000 - 8,000 Hz and 2,000 - 4,000 Hz were sow farrow, piglet voice and piglet's cry, respectively. The sound of the piglet being pressed was as similar as the sow farrow voice, which also falls at 2,000 - 4,000 Hz, with a similar frequency spectrum. This research might be to use image recognition to assist artificial intelligent judgment and to develop a smart farrowing care system in the future.

Key words: Pig farm, Robot, Artificial intelligence, Parturition.

(1) Contribution No. 2765 from Taiwan Livestock Research Institute (TLRI), Ministry of Agriculture (MOA).

(2) Northern Region Branch, MOA-TLRI, Beidow, Changhua 52149, Taiwan, R. O. C.

(3) Forage Crops Division, MOA-TLRI, HsinHua, Tainan 71246, Taiwan, R. O. C.

(4) Technical Service Division, MOA-TLRI, HsinHua, Tainan 71246, Taiwan, R. O. C.

(5) Microsystem Center of the Industrial Technology Research Institute.

(6) Corresponding author, E-mail: laiuyi@mail.tlri.gov.tw