Unveiling the Potential of Audio Classification for Poultry Health Diagnosis: A Deep Learning Approach

Kadam Shivanjali Ambadas, Madhuranthakam S Kavya Sree,
R. Sujatha[0000-0002-1993- 7544]*
School of Computer Science Engineering and Information Systems, Vellore Institute of Technology, Vellore, India
r.sujatha@vit.ac.in

Abstract
Audio classification has become a notable area of interest, particularly for its promising uses in the domain of poultry farming, where early detection of health issues is crucial. However, there is a lack of research on using deep learning techniques for this purpose. This study aims to fill this gap by exploring deep learning's potential for detecting various health issues in poultry. We propose a machine learning approach using audio data, collected from healthy and unhealthy birds, including vocalizations, breathing patterns, and movement noises. After preprocessing and feature extraction, machine learning models were trained to classify audio signals into healthy and diseased categories. The evaluation metrics, including accuracy of 0.9423, precision of 0.88, recall of 0.80, and F1-score of 0.8380, were used to demonstrate promising results. This research contributes to developing non-invasive and cost-effective methods for early disease detection in poultry, enhancing animal welfare and industry sustainability.

Keywords: Poultry farming, deep learning, artificial intelligence, disease detection, animal health, audio classification

1 Introduction:

Raising poultry is crucial for ensuring food security worldwide and economic stability, meeting the rising demand for meat and eggs worldwide. However, ensuring the health of poultry flocks is essential for sustainable and profitable operations within the industry. Diseases among poultry pose significant challenges, leading to economic setbacks through reduced productivity, higher mortality rates, and the need for costly treatments.

Traditionally, disease detection in poultry has heavily relied on visual inspections, physical exams, and laboratory analyses of blood or tissue samples. While effective, these methods are often invasive, require specialized equipment, and depend on the expertise of trained personnel, making them impractical for routine monitoring in large-scale operations.

In recent times, the development of new technologies has accumulated increasing attention. Advancements in sensor technology, signal processing, and machine learning have opened new avenues for innovative approaches to poultry health surveillance.
One promising avenue involves analyzing audio data to identify subtle changes in poultry vocalizations and behaviors indicative of various disease states. Healthy and diseased poultry emit distinct sounds characterized by differences in frequency, amplitude, duration, and temporal patterns.

In the realm of this research endeavor, we put forth a methodology grounded in machine learning techniques, aimed at categorizing the well-being of poultry through the analysis of audio recordings. By collecting audio recordings from both healthy and diseased poultry populations and training machine learning models to classify these signals, the aim is to develop a non-invasive and automated system for early disease detection on poultry farms.

Such research endeavors hold immense promise for revolutionizing poultry health monitoring practices, thereby enhancing animal welfare, reducing production losses, and fortifying the sustainability of the poultry industry.

2 Literature Review

The exploration of audio classification for poultry health diagnosis within the realm of deep learning builds upon a rich foundation of research in both animal health monitoring and machine learning applications The sounds made by animals provide a rich source of data regarding their health, feelings, and conduct[1].In the poultry industry, conventional approaches to monitoring the well-being of birds have heavily depended on visual assessments and manual observations, which can be arduous tasks, susceptible to personal interpretations, and frequently susceptible to inaccuracies arising from human limitations. The advancement of precision livestock farming has led to a surge in research utilizing image and sound technologies for animal monitoring. These non-invasive PLF techniques capture visual and auditory data, enabling the observation and assessment of both individual and collective animal behaviours, as well as environmental conditions and overall animal welfare [2].Their work underscores the importance of non-invasive methods, such as audio analysis, in providing continuous and real-time monitoring of animal health status.

The sounds made by poultry are a significant measure for measuring their welfare, offering insights into their nutrition, development, and overall health. Moreover, research has explored the aural characteristics of these vocal signals as a means to monitor the health of poultry [3].Additionally, delved into acoustic features of vocalization signals for poultry health monitoring, providing insights into the distinct patterns associated with different health states [4]. Their research results highlighted the promising capability of machine learning algorithms to unravel intricate vocalization data, thereby facilitating disease identification and diagnosis.

Investigated the use of audio technology and advanced machine learning techniques for the automated identification of Newcastle disease [5].The investigation validated the viability of harnessing deep learning models to precisely pinpoint particular ailments through the analysis of vocalization cues. This scholarly endeavour furnished invaluable perspectives into the prospective utilization of cutting-edge machine learning methodologies for diagnosing diseases affecting poultry populations. Furthermore, Analysing acoustic signals to discern and categorize stress levels in laying hens is an essential research effort, highlighting the ability of audio signals to reflect the emotional and physiological states of poultry [6].The work emphasized the importance of early stress detection for effective disease prevention and management [7].

In summary, the literature surrounding audio classification for poultry health diagnosis showcases a growing body of research that underscores the potential of machine learning techniques, particularly deep learning, in leveraging vocalization signals for disease detection and health assessment.

3 Proposed Work

This study proposes an innovative approach to monitoring poultry health through audio classification using deep learning methodologies. The principal objective revolves around the development of an automated framework adept at discerning and classifying diverse poultry-related maladies and distress indicators through the analysis of audio recordings. The methodological approach encompasses a sequence of critical phases: collecting data, pre-processing, feature extraction, model training, and evaluation.

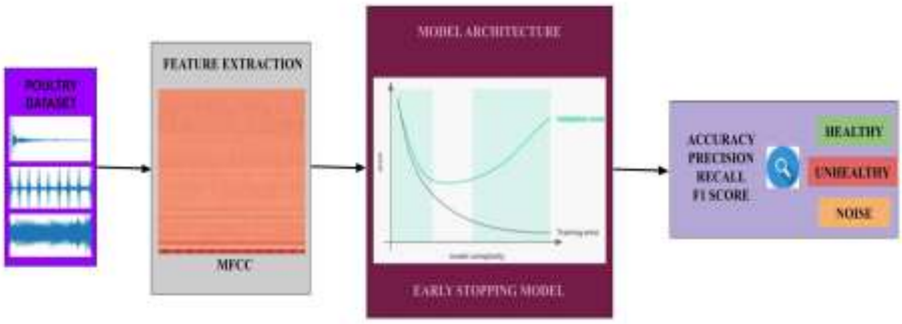

Fig. 1. Working Model

Figure 1 represents the working model. These are a few major steps in preprocessing then the model efficiency of the model is classified based on measures such as accuracy, precision, recall, and F1 score.

3.1 Data Collection

The preliminary stage entails gathering a diverse array of poultry audio recordings from commercial poultry farms and research institutions, encompassing a spectrum of health conditions like normal, healthy birds, alongside those displaying symptoms of prevalent diseases The collection comprises 346 audio recordings, categorized into three sections: 'Healthy' with 139 entries, 'Noisy' encompassing 86, and 'Unhealthy' containing 121. These recordings vary in duration from 5 to 60 seconds and are saved as .wav files. The 'Unhealthy' category consists of sounds such as coughs, snores, and rales of chickens, whereas the 'Noisy' category is made up of ambient sounds and noises produced by the activities of the poultry birds [8].

https://data.mendeley.com/datasets/zp4nf2dxbh/1

3.2 Preprocessing and Feature Extraction

Preprocessing procedures like noise reduction, normalization, and segmentation are implemented to enhance the audio data quality. Data preprocessing involves recognizing, defining, and detailing the problems associated with data, along with applying an informed strategy to resolve these problems, thereby enhancing the data's dependability for machine learning research [9]. In our research, we utilized Mel-Frequency Cepstral Coefficients (MFCCs) as the feature representation for the poultry audio signals. Mel Frequency Cepstral Coefficients (MFCCs) are commonly employed in audio signal processing [10] due to their efficiency in encapsulating the signal's spectral properties in a compact and perceptually meaningful way [11]. The choice of 40 MFCC coefficients is a common practice, as it provides a good trade-off between capturing sufficient spectral information and avoiding overfitting. Techniques for extracting features, like Mel-Frequency Cepstral Coefficients (MFCCs) [12], spectral centroid, zero-crossing rate, and energy, are employed to transform the audio signals into discernible features suitable for deep learning models [13]. Mel-Frequency Cepstral Coefficients (MFCCs) are [10] commonly employed in the extraction of features from audio signals and have proven effective in numerous sound classification attempts, such as the examination of animal sounds. The Mel-Frequency Cepstral Coefficient (MFCC) technique is widely used for extracting key features from audio signals, proving highly effective in various audio classification tasks, including animal vocalization analysis. MFCCs are engineered to mimic human auditory perception by incorporating the Mel scale, a nonlinear frequency scale aligned with how humans perceive pitch. Calculating MFCCs involves pre-emphasis, framing and windowing, Fourier transform, Mel filter bank application, logarithmic operation, and discrete cosine transform [10]. This multi-step process aims to capture salient audio features while emulating how the human hearing system reacts to [10] sound. In the context of poultry audio classification, MFCCs have been successfully employed to extract discriminative features from poultry vocalizations, enabling the identification of healthy, unhealthy and noisy audio.

3.3 Model Training and Evaluation

The fundamental methodology involves the systematic training of sophisticated deep learning architectures, containing Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), to enhance predictive modeling capabilities[14] to classify the audio recordings into distinct categories representing various poultry health conditions The data is segmented into different sets for training, validation, and testing, which are utilized for the instruction of the model and the assessment of its efficacy. In the training phase, the method of early stopping is utilized [15]. This approach keeps track of the model's progress on a validation dataset and halts the training when there are no further improvements observed, thereby preventing the risk of the model becoming too specific to the training data. Once training has been concluded with the implementation of early stopping, the model undergoes an evaluation phase on a reserved test set [16].

3.4 Entire Process

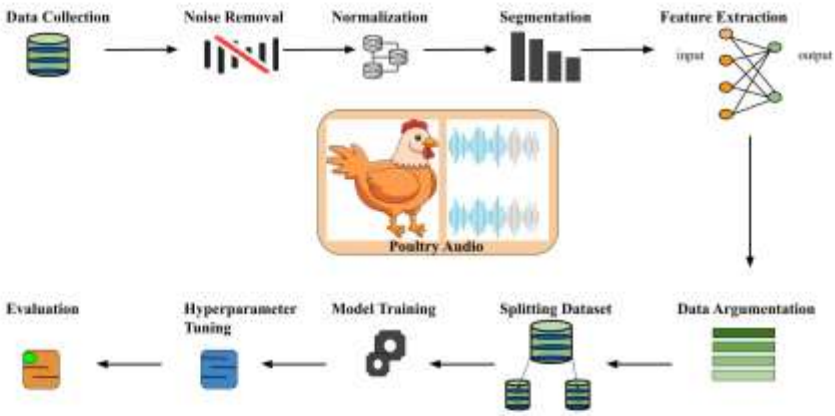

Fig. 2. Preprocessing Steps

Figure 2 represents the preprocessing steps that were implemented for the project flow.

Data Collection: A diverse dataset of poultry audio recordings is collected from various farms and research facilities. It includes healthy birds and those with common diseases like avian influenza or Newcastle disease, capturing different vocalizations such as distress calls and feeding sounds.

Noise Removal: Background noise in the audio recordings is eliminated using noise removal filters, preserving important features related to poultry vocalizations.

Normalization: To ensure dataset consistency, amplitude levels of audio signals are normalized to a standard scale, preventing biases in the model due to volume variations.

Segmentation: Audio recordings are segmented to isolate individual calls or vocalizations. This is crucial for extracting specific features related to different poultry vocalizations like distress calls or feeding sounds.

Feature Extraction: Techniques like Mel-Frequency Cepstral Coefficients (MFCCs) are used to transform audio signals into numerical representations suitable for deep learning.

Data Augmentation: When the dataset is limited, techniques such as pitch shifting and time stretching are used to create variations of the original recordings, increasing dataset diversity.

Splitting Dataset: The collected data is partitioned into three distinct subsets: one for training purposes, another for validation, and a third for testing the model's performance. Training is for model training, validation for hyperparameter fine-tuning and avoiding overfitting, and testing for evaluating model performance.

Model Training: Models learn to identify patterns in audio data indicative of different poultry health conditions.

Hyperparameter Tuning: Hyperparameters like learning rate and batch size are optimized during training to improve model performance and generalization [17].

Evaluation: The assessment of the trained model is carried out through various statistical measures such as accuracy, precision, recall, F1-score, and the confusion matrix to calculate its predictive accuracy [18].

4 Experimental Results and Discussions

4.1 Evaluation Metrics

Accuracy

The accuracy measure calculates the proportion of correct predictions generated by the model [19]. This measure serves as an indicator of the model's overall predictive proficiency, and its mathematical formulation is represented by Equation 1.

$$\frac{TruePositive+TrueNegative}{TruePositive+FalsePositive+TrueNegative+FalseNegative} \tag{1}$$

Precision

Precision is a measure that evaluates the proportion of correct positive predictions out of the total positive predictions made by the model [20]. It reflects the model's capability to identify and classify only the relevant instances as positive cases. The mathematical formulation of precision is elucidated by Equation 2.

$$\frac{True\ Positive}{TruePositive + FalsePositive} \tag{2}$$

Recall

Recall, often termed sensitivity, is the metric that determines the fraction of actual positive cases that the model correctly identifies. It measures the model's capacity to detect all instances that are indeed positive [21]. Shown in Equation 3.

$$\frac{TruePositive}{TruePositive + FalseNegative} \tag{3}$$

F1 Score

The F1 Score, as defined in Equation 4, is the harmonic mean of precision and recall. It offers a composite metric that equally weighs both precision and recall for a unified assessment of the model's performance [22].

$$\frac{2 * Precision * Recall}{Precision + Recall} \tag{4}$$

The successful classification and diagnosis of poultry diseases through vocalization analysis offers practical implications for poultry farmers and veterinarians, providing a valuable tool for routine health monitoring in poultry farms. By regularly analyzing vocalizations, farmers can detect subtle changes indicative of health issues, enabling proactive measures to maintain flock health.

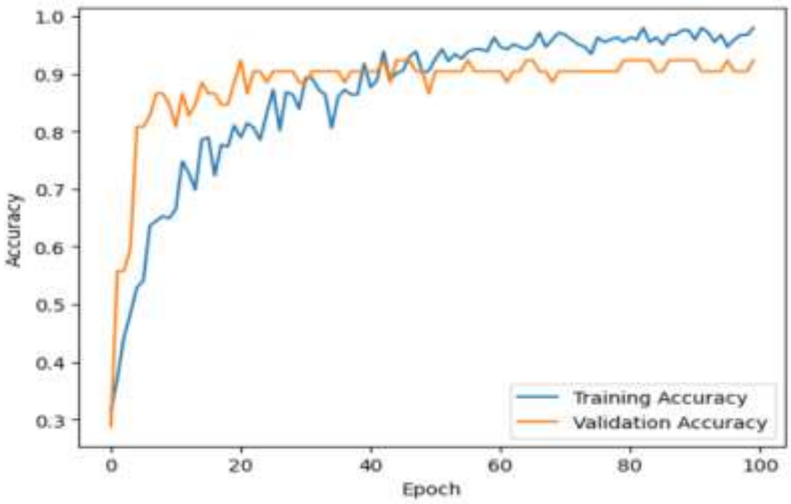

Fig. 3. Training and validation accuracy of the model

Figure 3 represents the training and validation accuracy of the model. These line plots provide valuable insights into the model's learning behavior and help determine when to stop training or adjust hyper parameters to achieve optimal performance.

Confusion matrix:

This offers a delicate perspective on the model's effectiveness, illustrating the count of instances that were accurately or inaccurately categorized.

	PREDICTED NEGATIVE	PREDICTED POSITIVE
ACTUAL NEGATIVE	TRUE NEGATIVE (TN) 41	FALSE POSITIVE (FP) 1
ACTUAL POSITIVE	FALSE NEGATIVE (FN) 2	TRUE POSITIVE (TP) 8

Fig. 4. Confusion Matrix

Figure 4 represents the confusion matrix for the testing dataset. The matrix is structured into four quadrants, each corresponding to the outcomes of the model's predictions

5 Conclusion and future work

The deep learning model for poultry health diagnosis through audio classification represents a significant advancement in automated disease detection. Its high accuracy of 0.9423, precision of 0.88, recall of 0.80, and F1 score of 0.8380 which demonstrate effectiveness in classifying poultry diseases based on vocalization signals. The proposed methodology presents a more impartial and streamlined approach to overseeing health-related concerns, bearing significant implications for disease mitigation strategies, enhancement of farm productivity, and ultimately, bolstering profitability within the agricultural sector. Integrating this model into farm management systems enables real-time monitoring, informed decision-making, and advancements in precision livestock farming. Addressing challenges such as dataset expansion, data privacy, and animal welfare considerations is crucial for further advancements in poultry health management through technology.

Continuous improvement and adaptation of the system are essential, considering technological advancements and user feedback. Conducting comprehensive evaluations, including testing and user feedback, ensures the system's performance and reliability. Future directions involve expanding the dataset with diverse poultry diseases, enhancing model interpretability, real-world implementation through field trials, collaboration with veterinarians, and investigating long-term monitoring capabilities.

References

1. Briefer, E. F. (2012). Vocal expression of emotions in mammals: mechanisms of production and evidence. Journal of Zoology, 288(1), 1-20.
2. Aydin, A. (2017). Using 3D vision camera system to automatically assess the level of inactivity in broiler chickens. Computers and Electronics in Agriculture, 135, 4-10.
3. Fontana, I., Tullo, E., Scrase, A., & Butterworth, A. (2016). Vocalisation sound pattern identification in young broiler chickens. Animal, 10(9), 1567-1574.
4. Ferrari, S., Silva, M., Guarino, M., Aerts, J. M., & Berckmans, D. (2008). Cough sound analysis to identify respiratory infections in pigs. Computers and Electronics in Agriculture, 64(2), 318-325.
5. Carpentier, L., Vranken, E., Berckmans, D., Paeshuyse, J., & Norton, T. (2019). Development of sound-based poultry health monitoring tool for automated sneeze detection. Computers and electronics in agriculture, 162, 573-581.
6. Lee, J., Noh, B., Jang, S., Park, D., Chung, Y., & Chang, H. H. (2015). Stress Detection and Classification of Laying Hens by Sound Analysis. Asian-Australasian Journal of Animal Sciences, 28(4), 592–598.
7. Astill, J., Dara, R. A., Fraser, E. D., & Sharif, S. (2018). Detecting and predicting emerging disease in poultry with the implementation of new technologies and big data: A focus on avian influenza virus. Frontiers in veterinary science, 5, 263.
8. Adebayo, S., Aworinde, H. O., Akinwunmi, A. O., Alabi, O. M., Ayandiji, A., Sakpere, A. B., Adeyemo, A., Oyebamiji, A. K., Olaide, O., & Kizito, E. (2023). Enhancing poultry health management through machine learning-based analysis of vocalization signals dataset. Data in Brief, 50, 109528.

9. Zhang, S., Zhang, C., & Yang, Q. (2003). Data preparation for data mining. Applied artificial intelligence, 17(5-6), 375-381.
10. Mehrish, A., Majumder, N., Bharadwaj, R., Mihalcea, R., & Poria, S. (2023). A review of deep learning techniques for speech processing. Information Fusion, 99, 101869.
11. Umapathy, K., Ghoraani, B., & Krishnan, S. (2010). Audio signal processing using time-frequency approaches: coding, classification, fingerprinting, and watermarking. EURASIP Journal on Advances in Signal Processing, 2010, 1-28.
12. Muda, L., Begam, M., & Elamvazuthi, I. (2010). Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques. Journal of Computing, 2(3).
13. Gold, B., Morgan, N., & Ellis, D. (2011). Speech and audio signal processing: processing and perception of speech and music. John Wiley & Sons.
14. Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. Journal of Big Data, 8, Article 53.
15. Prechelt, L. (2002). Early stopping-but when?. In Neural Networks: Tricks of the trade (pp. 55-69). Berlin, Heidelberg: Springer Berlin Heidelberg.
16. Hossin, M., & Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. International journal of data mining & knowledge management process, 5(2), 1.
17. Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. Journal of machine learning research, 13(2).
18. Boyko, N. (2023). Evaluating Binary Classification Algorithms on Data Lakes Using Machine Learning. Revue d'Intelligence Artificielle, 37(6), 1423-1434.
19. Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). The elements of statistical learning: data mining, inference, and prediction (Vol. 2, pp. 1-758). New York: springer..
20. Sharma, P. (2020, September 1). Precision and recall in machine learning. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2020/09/precision-recall-machine-learning/
21. Baghirov, E. (2024). A Comprehensive Investigation into Robust Malware Detection with Explainable AI. Cyber Security and Applications, 100072.
22. Powers, D. M. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061*.