Automated Poultry Health Monitoring through Acoustic Analysis Using Convolutional Neural Networks

H. M. Rohini
Department of Electronics & Communication Engineering, RYMEC Ballari, Karnataka, India | Department of Electronics & Communication Engineering, Central University of Karnataka, Kalaburgi, Karnataka, India
rohinihalakatti@rymec.in (corresponding author)

S. Prabhavathi
Department of Electronics & Communication Engineering, RYMEC Ballari, Karnataka, India
prabkir@rymce.in, prabkir@gmail.com

Received: 21 April 2025 | Revised: 8 May 2025, 25 May 2025, and 8 June 2025 | Accepted: 21 June 2025

 DOI: https://doi.org/10.48084/etasr.11622

ABSTRACT

Improving animal welfare and reducing losses in poultry breeding and production systems hinge on the early detection and warning of contagious diseases among chickens. Traditional methods for controlling and diagnosing poultry diseases often fall short, leading to significant mortality and decreased output. This study presents an automated poultry health management algorithm based on Convolution Neural Networks (CNNs) to identify healthy and unhealthy chickens using acoustic analysis of their vocalizations. The proposed approach leverages Mel-Frequency Cepstral Coefficients (MFCC) for feature extraction from audio signals of chickens that exhibit respiratory diseases. The CNN model, which comprises convolution layers, dropout layers, and batch normalization, was trained and evaluated on a dataset of 346 audio signals collected from poultry farms. The results demonstrate high accuracy (94.59%), precision (96%), recall (96%), and F1-score (96%) in classifying healthy and unhealthy chicken sounds, outperforming previous methods. This study underscores the potential of voice-based diagnostic tools in poultry health management, offering prospects for early intervention and enhanced health outcomes.

Keywords-acoustic waveform; CNN; Mel-Frequency Cepstral Coefficients (MFCC); poultry vocalization

I. INTRODUCTION

As cattle and poultry populations are vulnerable to contagious diseases that can have serious financial repercussions for agriculture, voice-based disease detection offers a potential way to keep an eye on their health [1]. By incorporating sound-based monitoring systems into livestock management procedures, farmers and veterinarians can maximize production efficiency, protect animal welfare, and prevent disease outbreaks proactively [2]. Advanced tools for analyzing cattle and poultry vocalizations have been developed as a result of technological breakthroughs, especially in the fields of bioacoustics and machine learning [3]. These instruments enable scientists to identify patterns related to certain diseases and extract important information from complex noises made by animals [4]. The non-invasive feature of voice-based disease identification is beneficial and makes it easier to remotely monitor animal populations, enabling prompt action and early epidemic detection [5]. Researchers can collect a lot of data by carefully placing sound recording equipment, which can give insight into the dynamics of disease and its transmission patterns [6].

Voice-based surveillance also helps detect and control zoonotic infections, which can spread from animals to people. Researchers can implement preventive measures and reduce the risk of spillover occurrences and possible pandemics by identifying and tracking these diseases in animal populations [7]. The subject of vocalization-based animal disease detection is still in its infancy [8], but it has great potential to revolutionize public health initiatives, wildlife conservation, and veterinary treatment [9]. Furthermore, studying animal vocalizations provides important insights into animal behavior and wellbeing in addition to its use in disease diagnosis [10]. By studying variations in speech patterns, scientists can gain a deeper understanding of how animals interact, communicate, and react to their environment [11].

This study aimed to implement a poultry health management algorithm to identify healthy and unhealthy chickens in poultry farms, to help with early diagnosis of respiratory diseases, isolate chickens, and control the spread of infectious diseases. Table I shows details of the dataset used in this study. Figure 1 shows a flow chart of the proposed automated poultry health monitoring framework.

TABLE I. DATA SPECIFICATION TABLE

Data type	Audio Signal
Data collection	100 day-old poultry birds were purchased and divided into two groups at the experimental site, the poultry research farm at Bowen University. For respiratory diseases, the first group received treatment, while the second group did not. After that, the birds were separated and caged in a monitored environment. To eliminate extraneous sounds and background noise that might affect the analysis, microphones were set a reasonable distance from the birds. The data were gathered using 24-bit samples at 96 kHz. Audio data were collected for 65 days, three times per day (morning, afternoon, and night). During this time, the birds had constant access to food and water. After 30 days, the untreated group started to sound sick with respiratory issues. This information was also noted as unhealthy. Chicken' audio signals were recorded, saved in MA4, and then converted to the wav format.
Data format	Raw audio signals
Data description	346 audio signal files, organized into three folders: healthy, noisy, and sick. There are 139 audio files in the healthy folder, 86 in the noise folder, and 121 in the unhealthy folder.
Data source location	• Institution: Bowen University, Poultry Research Farm, Bowen University Commercial Farm • City/Town/Region: Iwo • Country: Nigeria
Data accessibility	[12]
Data article	[3]

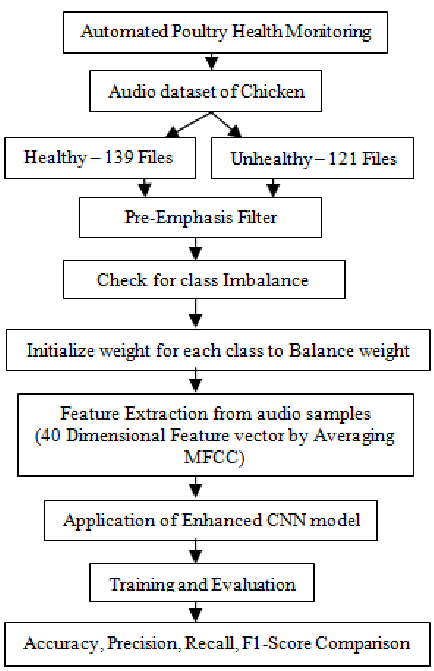

Fig. 1. Flow chart of the proposed automated poultry health monitoring method.

II. MATERIALS AND METHODS

This study used the dataset in [3, 12], which consists of 346 audio signal files in wav format. 139 audio files are acoustic emissions of healthy chickens, and 121 audio files are acoustic emissions of unhealthy chickens experiencing respiratory disease. Acoustic emissions were captured for the two groups of chickens at the experimental site of Bowen University in a monitored environment to collect a dataset for early detection of respiratory diseases. As part of the preprocessing of the sound signals, the sound samples were framed and filtered. A moving Hamming window was then used to frame longer period non-stationary sound samples, producing a stationary signal that lasted 10 to 30 seconds.

Figures 2 and 3 show the acoustic waveforms of healthy and unhealthy chicken sound signals, respectively.

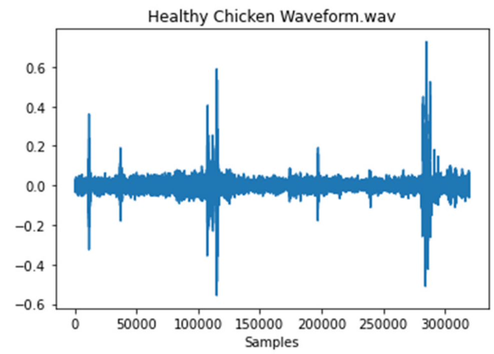

Fig. 2. Acoustic waveform of a healthy chicken sound signal.

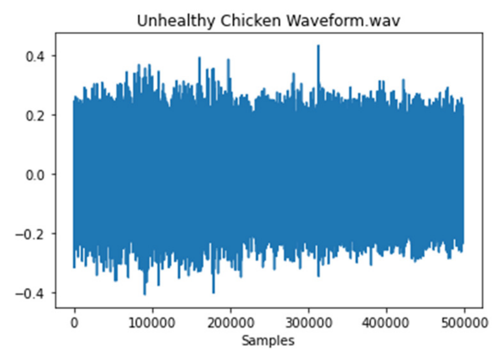

Fig. 3. Acoustic waveform of an unhealthy chicken sound signal.

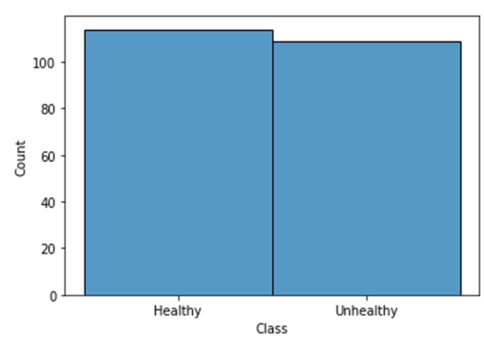

Fig. 4. Class imbalance for the number of healthy and unhealthy audio samples considered.

The weights for two classes were calculated as follows. Let N be the total number of training samples, N_x be the number of samples in class x, where $x \in$ {Healthy, Unhealthy}. For each class x, the weight is:

$$W_x = 1 - N_x/N, \text{ for } x \in \{\text{Healthy, Unhealthy}\} \quad (1)$$

where W_x is the weight assigned to class x, N_x is the number of samples in class x, and N is the total number of training samples. The weight for the healthy chicken audio samples was calculated as 0.4888, and the weight for the unhealthy chicken audio samples was calculated as 0.5112, indicating a small imbalance. Figure 4 shows a histogram that depicts class imbalance for the number of healthy and unhealthy audio signals considered. The Torchaudio transform resample ensured that all samples had an equal sample rate of 22050 Hz. The Torchaudio transform MelSpectrogram converted the audio files to Mel spectrograms, which are intuitively the equivalent of converting an audio file to an image. After this conversion, audio classification was modified to an image classification problem. As the dataset had audio files of varying sizes for image classification, a Torchvision transform helped to resize every Mel Spectrogram to a fixed shape (256, 256). For the feature extraction, the Mel-Frequency Cepstral Coefficients (MFCC) algorithm was used [13, 14].

Given that people are more adept at identifying distinctions in lower frequencies than in higher frequencies, a Mel spectrogram converted frequencies to the Mel scale. Figures 5 and 6 show the spectra of healthy and unhealthy chicken sounds, respectively. Figures 7 and 8 show the Mel-Frequency Cepstral Coefficients for healthy and unhealthy chicken sounds, respectively.

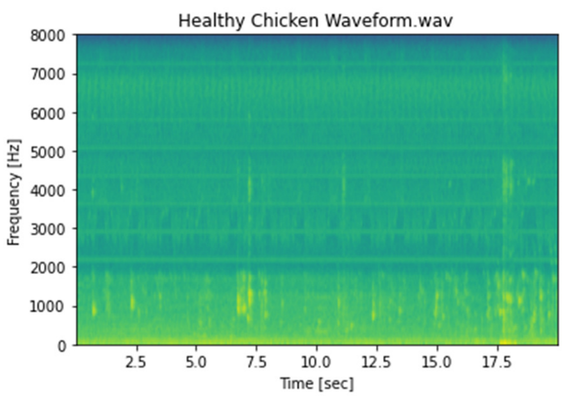

Fig. 5. Spectrum of healthy chicken sound signal.

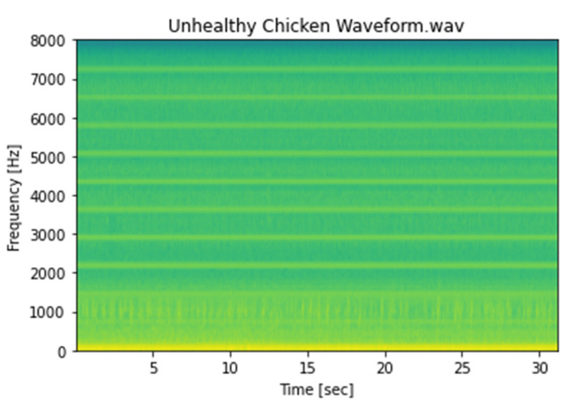

Fig. 6. Spectrum of unhealthy chicken sound signal.

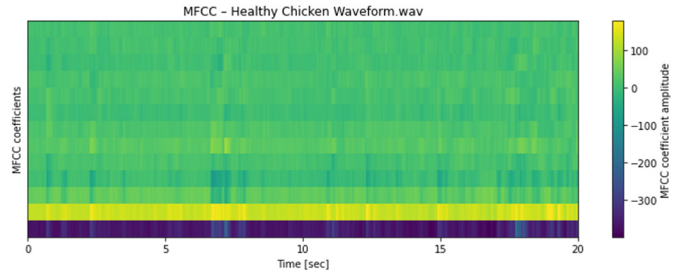

Fig. 7. MFCC of healthy chicken sound signal.

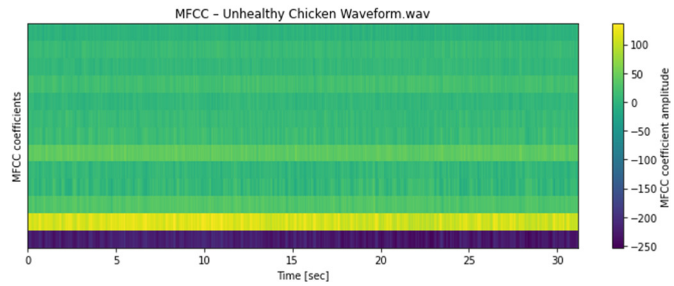

Fig. 8. MFCC of unhealthy chicken sound signal.

A. *Convolution Neural Network Model*

The CNN/ConvNet class of deep neural networks is most frequently used in deep learning for visual imagery analysis. Convolution is a mathematical operation on two functions that yields a third one that expresses how the shape of one is altered by the other. When the accuracy of the training dataset is higher than the testing accuracy, machine learning models are said to be overfitting or to have a high variance. When a model has a low error in the training set and a higher error in the testing set, overfitting is evident in terms of loss. The underlying model was altered to incorporate dropout layers, a batch normalization layer, and early stopping in the training loop to reduce overfitting. Algorithm 1 shows the model modified to counteract overfitting.

```
model=nn.Sequential(
 nn.Conv2d(in_channels=1, out_channels=32,
  kernel_size=3, stride=2),
 nn.BatchNorm2d(32),
 nn.ReLU(),
 nn.Conv2d(in_channels=32,
  out_channels=64, kernel_size=3,
  stride=2),
 nn.BatchNorm2d(64),
 nn.ReLU(),
 nn.Flatten(),
 nn.Dropout(0.4),
 nn.Linear(64*63*63, 512),
 nn.BatchNorm1d(512),
 nn.ReLU(),
 nn.Dropout(0.4),
 nn.Linear(512,2))
```

III. RESULTS AND DISCUSSION

Out of 139 audio files of acoustic emissions of healthy chickens, 114 audio files were considered for training the model, and 25 audio files were kept for testing. Out of 121 audio files of acoustic emissions of unhealthy chickens, 109 audio files were considered for training the model, and 12 audio files were kept for testing. Table II and Figure 9 show Accuracy, Precision, Recall, and F1-score for both the training and testing of the model. Table III shows the confusion matrix for training, and Table IV shows the confusion matrix for testing.

TABLE II. ACCURACY, PRECISION, RECALL, AND F1-SCORE FOR TRAINING AND TESTING

Phase	Accuracy	Precision	Recall	F1-Score
Training	99.55%	100%	99.13%	99.56%
Testing	94.59%	96%	96%	96%

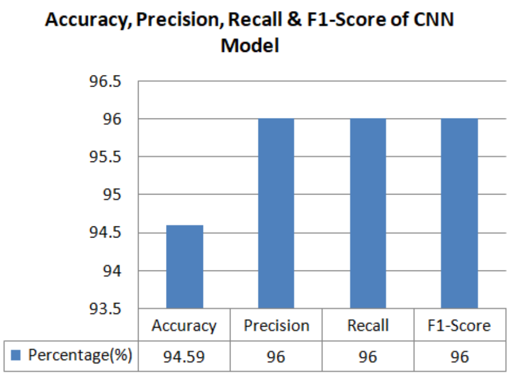

Fig. 9. Accuracy, Precision, Recall, and F1-Score of the CNN model.

TABLE III. CONFUSION MATRIX FOR TRAINING

Training	Healthy	Unhealthy
Healthy	24 (TP)	1 (FP)
Unhealthy	1 (FN)	11 (TN)

TABLE IV. CONFUSION MATRIX FOR TESTING

Testing	Healthy	Unhealthy
Healthy	114 (TP)	0 (FP)
Unhealthy	1 (FN)	108 (TN)

Figure 10 shows the test and training losses of the CNN. The training loss indicates how far the predicted values are from the actual values (targets) for the training set. The goal is to minimize training loss throughout training, which is achieved as shown in Figure 10. Lower training loss means that the model is learning well from the training data. The test loss is calculated using the test data (which is not used during training) and evaluates how well the model generalizes to unseen data. A good model should have low training and test losses, indicating that it has learned the underlying patterns and is not overfitting or underfitting.

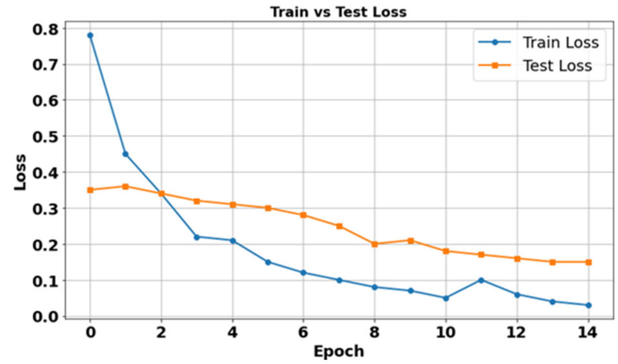

Fig. 10. Test and training losses of the CNN model.

In Epoch 15, both test and training loss curves are almost parallel to the x-axis, which means that the model is fully trained and further training will not provide improvements. Thus, training stopped at Epoch 15. The gap between test and training losses gradually decreases and reaches a minimum gap by Epoch 15, which resembles a good fit.

Table V shows a comparison of the proposed method with previous ones. The results show that the proposed algorithm performs better. Technological advances, particularly in fields such as bioacoustics and machine learning, have facilitated the collection, analysis, and interpretation of vocalization data. Previous studies have highlighted the efficacy of voice-based diagnostic tools in identifying diseases in poultry, offering prospects for early intervention and enhanced health outcomes. Looking ahead, voice-based disease detection is poised to revolutionize poultry health monitoring. By leveraging the distinct vocal signatures associated with different diseases, researchers aspire to develop precise, cost-effective, and scalable tools for disease detection and management, thereby contributing to the well-being of both animals and humans.

TABLE V. ACCURACY COMPARISON WITH PREVIOUS STUDIES

Ref.	Method	Result
[1]	VGG Net-16 and ResNet-50	VGGNet-16: 74.1% Accuracy, ResNet-50: 66.91% Accuracy
[2]	3D vision camera system for activity level	93% classification accuracy
[5]	Sound signal analysis for avian influenza	Accuracy rate ranged between 84% and 90%.
[7]	Thermography + AI	92.85 Average Accuracy based on different Hours of response.
[8]	Extreme learning machine and SVM	ELM achieved 88% precision and 81% recall. SVM achieved 86% precision and 85% recall.
[9]	ResNet50	75.6% Accuracy
[10]	Deep CNN for image-based disease detection	Accuracy of 93.67% with a fully trained CNN.
Proposed method	CNN-based classifier trained on MFCC features extracted from vocalizations of healthy and unhealthy chickens Audio (Preprocessing involved noise reduction, framing, and normalization)	Accuracy-94.59%, Precision-96%, Recall-96%, F1-Score-96%.

IV. CONCLUSION AND FUTURE WORK

The proposed Automated Poultry Health Management through acoustic analysis using CNN addresses the critical issue of early disease detection in poultry by leveraging audio-based classification techniques. While traditional methods rely on visual or invasive biological assessments, this research fills a notable gap by focusing on bioacoustic signals using deep learning techniques. The method employs a CNN-based classifier trained on MFCC features extracted from vocalizations of healthy and unhealthy chickens. Audio preprocessing included noise reduction and time-frequency transformation through spectrograms. The proposed model achieved high performance with 94.59% accuracy, 96% precision, 96% recall, and an F1-score of 96%, demonstrating its potential as a non-invasive, real-time monitoring tool.

Compared to previous studies, the proposed method aligns with state-of-the-art systems in terms of architecture (e.g., VGG, ResNet) but stands out by integrating lightweight preprocessing, end-to-end MFCC-based feature pipelines, and high precision in audio-only settings. The system offers practical value in real-world poultry farms where visual monitoring is infeasible, thereby supporting scalable and cost-effective disease surveillance. However, the proposed method was tested on a small dataset due to the difficulty of collecting data. In the future, data augmentation techniques can be used to increase the amount of data to effectively train the system, and ensemble machine learning algorithms can be used to further improve classification performance.

ACKNOWLEDGMENT

The authors acknowledge the work of the authors in [3].

REFERENCES

[1] L. D. Quach, N. Pham-Quoc, D. C. Tran, and M. F. Hassan, "Identification of Chicken Diseases Using VGGNet and ResNet Models," in *Industrial Networks and Intelligent Systems*, 2020, pp. 259–269, https://doi.org/10.1007/978-3-030-63083-6_20.

[2] A. Aydin, "Using 3D vision camera system to automatically assess the level of inactivity in broiler chickens," *Computers and Electronics in Agriculture*, vol. 135, pp. 4–10, Apr. 2017, https://doi.org/10.1016/j.compag.2017.01.024.

[3] S. Adebayo *et al.*, "Enhancing poultry health management through machine learning-based analysis of vocalization signals dataset," *Data in Brief*, vol. 50, Oct. 2023, Art. no. 109528, https://doi.org/10.1016/j.dib.2023.109528.

[4] M. S. Uddin *et al.*, "A Novel Hybrid Deep Neural Network for Early Detection and Classification of Chicken Diseases," in *2023 5th International Conference on Sustainable Technologies for Industry 5.0 (STI)*, Dhaka, Bangladesh, Dec. 2023, pp. 1–6, https://doi.org/10.1109/sti59863.2023.10464991.

[5] J. Huang, W. Wang, and T. Zhang, "Method for detecting avian influenza disease of chickens based on sound analysis," *Biosystems Engineering*, vol. 180, pp. 16–24, Apr. 2019, https://doi.org/10.1016/j.biosystemseng.2019.01.015.

[6] K. Cuan, T. Zhang, J. Huang, C. Fang, and Y. Guan, "Detection of avian influenza-infected chickens based on a chicken sound convolutional neural network," *Computers and Electronics in Agriculture*, vol. 178, Nov. 2020, Art. no. 105688, https://doi.org/10.1016/j.compag.2020.105688.

[7] M. Sadeghi, A. Banakar, S. Minaei, M. Orooji, A. Shoushtari, and G. Li, "Early Detection of Avian Diseases Based on Thermography and Artificial Intelligence," *Animals*, vol. 13, no. 14, Jul. 2023, Art. no. 2348, https://doi.org/10.3390/ani13142348.

[8] M. Rizwan *et al.*, "Identifying rale sounds in chickens using audio signals for early disease detection in poultry," in *2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Washington, DC, USA, Dec. 2016, pp. 55–59, https://doi.org/10.1109/globalsip.2016.7905802.

[9] F. J. Bravo Sanchez, M. R. Hossain, N. B. English, and S. T. Moore, "Bioacoustic classification of avian calls from raw sound waveforms with an open-source deep learning architecture," *Scientific Reports*, vol. 11, no. 1, Aug. 2021, https://doi.org/10.1038/s41598-021-95076-6.

[10] H. Mbelwa, J. Mbelwa, and D. Machuve, "Deep Convolutional Neural Network for Chicken Diseases Detection," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 2, 2021, https://doi.org/10.14569/ijacsa.2021.0120295.

[11] G. M. S. Himel and Md. M. Islam, "A Smart Intelligence System for Hen Breed and Disease classification using Extra Tree classifier-based Ensemble Technique," *Journal of Electrical Systems and Information Technology*, vol. 12, no. 1, Feb. 2025, https://doi.org/10.1186/s43067-025-00191-3.

[12] S. Adebayo, "Poultry Vocalization Signal Dataset for Early Disease Detection." Mendeley, Jun. 12, 2023, https://doi.org/10.17632/ZP4NF2DXBH.1.

[13] V. L. Hardjanto and Wahyono, "Audio Enhancement for Gamelan Instrument Recognition using Spectral Subtraction," *Engineering, Technology & Applied Science Research*, vol. 15, no. 2, pp. 22042–22048, Apr. 2025, https://doi.org/10.48084/etasr.10181.

[14] T. H. Tran, P. A. Nguyen, L. A. Ngoc, D. T. Tran, and M. T. Pham, "Optimal CNN Model for Obstructive Sleep Apnea Detection using Particle Swarm Optimization," *Engineering, Technology & Applied Science Research*, vol. 15, no. 1, pp. 19553–19560, Feb. 2025, https://doi.org/10.48084/etasr.9154.