Data Article

Enhancing poultry health management through machine learning-based analysis of vocalization signals dataset

Segun Adebayo [b], Halleluyah O. Aworinde [a,*],
Akinwale O. Akinwunmi [a], Olufemi M. Alabi [b], Adebamiji Ayandiji [b],
Aderonke B. Sakpere [c], Adetoye Adeyemo [a], Abel K. Oyebamiji [b],
Oke Olaide [a], Echentama Kizito [a]

[a] College of Computing and Communication Studies, Bowen University, Iwo Nigeria
[b] College of Agriculture, Engineering and Science, Bowen University, Iwo Nigeria
[c] Computer Science Department, University of Ibadan, Ibadan Nigeria

ARTICLE INFO

ABSTRACT

Population expansion and rising consumer demand for nutrient-dense meals have both contributed to an increase in the consumption of animal protein worldwide. A significant portion of the meat and eggs used for human consumption come from the poultry industry. Early diagnosis and warning of infectious illnesses in poultry are crucial for enhancing animal welfare and minimizing losses in the breeding and production systems for poultry. On the other hand, insufficient techniques for early diagnosis as well as infectious disease control in poultry farms occasionally fail to stop declining productivity and even widespread death.

Individual physiological, physical, and behavioral symptoms in poultry, such as fever-induced increases in body temperature, abnormal vocalization due to respiratory conditions, and abnormal behavior due to pathogenic infections, frequently represent the health status of the animal. When birds have respiratory problems, they make strange noises like coughing and snoring. The work is geared towards compiling a dataset of chickens that were both healthy and unhealthy.

---

* Corresponding author.
  E-mail address: aworinde.halleluyah@bowen.edu.ng (H.O. Aworinde).

100 day-old poultry birds were purchased and split into two groups at the experimental site, the poultry research farm at Bowen University. For respiratory illnesses, the first group received treatment, whereas 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 away from the birds. The data was gathered using 24-bit samples at 96 kHz. For 65 days, three times per day (morning, afternoon, and night) of audio data were continually collected. Food and water are constantly provided to the birds during this time. During this time, the birds have 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 being unhealthy. Chickens' audio signals were recorded, saved in MA4, and afterwards converted to WAV format.

This dataset's creation is intended to aid in the design of smart technologies capable of early detection and monitoring of the status of birds in poultry farms in a continuous, noninvasive, and automated way.

Specification Table

Subject	Applied Machine Learning
Specific Subject Area	Signal Processing
Type of Data	Audio Signal
How data were acquired	100 day-old poultry birds were purchased and split into two groups at the experimental site, the poultry research farm at Bowen University. For respiratory illnesses, the first group received treatment, whereas 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 away from the birds. The data was gathered using 24-bit samples at 96 kHz. For 65 days, three times per day (morning, afternoon, and night) of audio data were continually collected. Food and water are constantly provided to the birds during this time. During this time, the birds have 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 being unhealthy. Chickens' audio signals were recorded, saved in MA4, and afterwards converted to WAV format.
Data format	Raw Audio Signal
Description of Data Collection	There are 346 audio signal files altogether in the dataset, and they are 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. Wav files are used to store the audio files.
Data source Location	• Institution: Bowen University, Poultry Research Farm, Bowen University Commercial Farm • City/Town/Region: Iwo • Country: Nigeria
Data Accessibility	Repository name: Mendeley Data Data identification number: doi: Direct URL to data: doi:10.17632/zp4nf2dxbh.2, https://dx.doi.org/10.17632/zp4nf2dxbh.1

1. Value of the Data

- Audio signal dataset of the chicken fecal have potential value in early detection mechanisms and diagnosis of poultry diseases due to high contagious rate of such diseases.
- Respiratory related diseases have been discovered to have damaging effects on the vocal tract of poultry and which thereby affect both their breathing behaviour and the sound they produced. Hence, early observation and classification for vocal data gotten from them is imperative.
- There is paucity of publicly available datasets of poultry vocalization signals for respiratory related diseases in poultry. Making annotated dataset gotten from this work available will promote feasibility, accessibility, interoperability, and reproducibility (FAIR Data) a possibility.
- The dataset is applicable in the development of computer vision-based models for audio signals, segmentation, and classification.
- The dataset will give more insights beyond human physical observation for accurate and precise detection of poultry diseases. It will aid agricultural extension agents in engaging poultry farmers in a more effective training and orientation.

2. Objective

Due to issues with stocking density and inadequate management brought on by the growth of the large-scale and intensive broiler business, the incidence of respiratory illnesses such as infectious bronchitis, Newcastle disease, and avian influenza has grown. When broilers have respiratory conditions, they will make strange noises like coughing and snoring.

Animals' vocalizations offer a wealth of information about their well-being, emotions, and behavior [1]. These have propelled many researchers to use various techniques for analyzing animal vocalizations in recent years [2].

Acoustic technology has been used to analyze growth rate in poultry birds [3,4], disease detection [5–7], sex of birds [8], poultry birds under heat or fear stress [9–11]. Despite the fact that these studies have demonstrated the value of using animal sound analysis as a tool for early disease, stress, and behavior detection in several animal species, unfortunately, many of these datasets are not publicly available on open source repositories [12]. This study aims at addressing this gap by providing a labeled auditory dataset of healthy and unhealthy chicken sounds for commonly reared poultry birds' species in Africa on an open-source platform.

3. Data Description

The dataset contains 346 audio signal files and grouped into three folders: healthy, noisy, and unhealthy. There are 139 audio files in the healthy folder, 86 in the noise folder, and 121 in the unhealthy folder. The audio signal file has a range of 5 s to 60 s length. The selected sound segments in the unhealthy folder include chicken cough, snore and rale sound. The Noise folder contain sound segments that include background noises (e.g., moving vehicle and human voices) and sound created by poultry bird activities (e.g., feeding, pecking one another). The files are stored in .wav audio format.

In order to investigate, the audio signal was analyzed in the time and frequency domains. The power spectra of the two signals (healthy and unhealthy) were compared in the frequency domain. While the normal sound has more frequency content exactly around the y-axis, the anomalous sound exhibits a noticeable spike of about 0.2 radians.
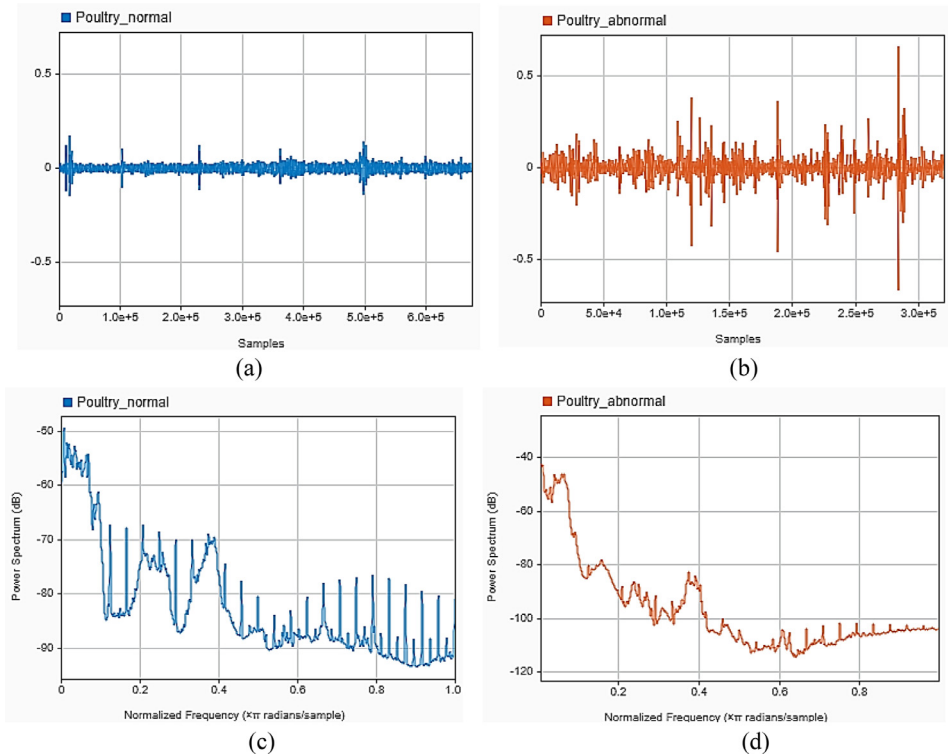
Fig. 1. Poultry Sound in time and frequency domain (a) Normal Poultry in time domain, (b) Abnormal Poultry Sound in time domain

4. Experimental Design, Materials and Methods

4.1. Field Data Collection

100 day-old poultry birds were purchased and split into two groups at the experimental site, the poultry research farm at Bowen University. For respiratory illnesses, the first group received treatment, whereas 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 away from the birds. The data was gathered using 24-bit samples at 96 kHz. For 65 days, three times per day (morning, afternoon, and night) of audio data were continually collected. Food and water are constantly provided to the birds during this time. During this time, the birds have access to food and water constantly. After 30 days, the untreated group started to sound sick with respiratory issues. This information was also noted as being unhealthy. Chickens' audio signals were recorded, saved in MA4, and afterwards converted to WAV format (Figs. 1,2).

4.2. Data Preprocessing

The sound samples were framed and filtered as part of the sound signal preprocessing. Longer period nonstationary sound samples were then framed by a shifting Hamming window to create a 10 to 30 s stationary signal.
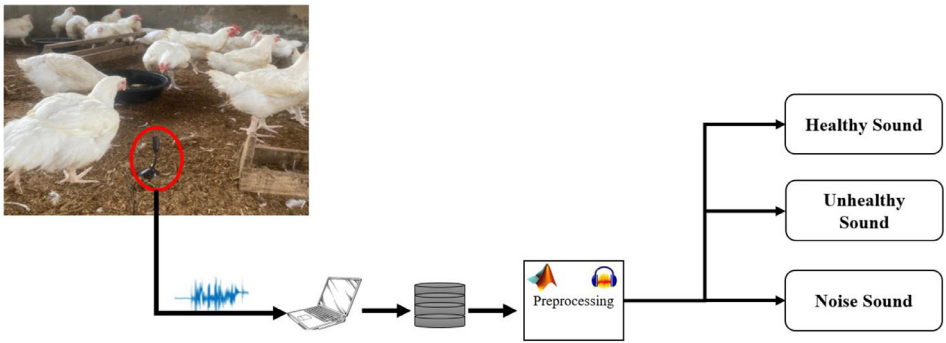
Fig. 2. Data acquisition experimental setup.

The Hann window [13] is expanded by the Hamming window in that it is a higher cosine window of the same type. The Hamming window is given as:

$$h(n) = \alpha + (1.0 - \alpha)\cos\left[\left(\frac{2\pi}{N}\right)\right] \tag{1}$$

with a matching form spectrum

$$H(\theta) = \alpha D(\omega) + \frac{(1.0 - \alpha)}{2}\left[D\left(\omega + \frac{2\pi}{N}\right)\right] \tag{2}$$

α enables the destructive sidelobe cancellation to be optimized. The window is known as the Hamming window and has the form when the value of is roughly 0.54.

$$H(\theta) = 0.54 + 0.46\cos\left[\left(\frac{2\pi}{N}\right)n\right], \ 0 \le n \le N \tag{3}$$

The window's size is L=N+1.

In order to improve the signal-to-noise ratios of the acoustic samples, filtering was done using the Kalman filter. This is a nonstationary, recursive filter that enables an estimate of the usable signal in noisy time series at each instant of time [14]. Two linear difference stochastic equations define the Kalman technique for digital filtering signal $\{x_k\}$ in steady state.

$$x_k = \mathcal{F}x_{k-1} + \mathcal{G}w_{k-1}$$

$$y_k = \mathcal{H}x_k + \nu_k \tag{4}$$

The first equation represents the signal generation process, whereas the second equation describes the signal measurement process.

The measurement noises are represented by the random variables w_k and ν_k, respectively. These variables can be assumed to be white, unrelated to one another, and to have normal probability distributions.

In the absence of a driving function or process noise, the matrix $\mathcal{F}$ links the state at time step 'k-1' to the state at step 'k'. Also, the matrix $\mathcal{H}$ relates the state x_k to the measurement y_k.

At step 'k', the Kalman filter makes an estimate of the process state and subsequently receives input in the form of (noisy) measurements.

Thus, two groups of Kalman filter equations can be distinguished, as shown in equations 5 and 6.

Equation 5 is the time update, while equation 6 is the measurement update:

$$\hat{x}_{k|k-1} = \mathcal{F}\hat{x}_{k-1|k-1}$$

$$\mathcal{P}_{k|k-1} = \mathcal{F}\mathcal{P}_{k-1|k-1}\mathcal{F}^T + \mathcal{Q} \tag{5}$$

$$\varkappa_k = \mathcal{P}_{k|k-1} \mathcal{H}_k^T \big[\mathcal{H}_k \mathcal{P}_{k|k-1} \mathcal{H}_k^T + \mathcal{R} \big]^{-1}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + \varkappa_k \big(y_k - \mathcal{H}_k \hat{x}_{k|k-1} \big)$$

$$\mathcal{P}_{k|k} = [\updownarrow - \varkappa_k \mathcal{H}_k] \mathcal{P}_{k|k-1} \tag{6}$$

Where matrix $\mathcal{P}_{k|k-1}$ is a priori estimate of error covariance.

Ethics Statements

The Bowen University Research Ethical Board examined and approved the animal study. The research was conducted under the required ethical guidelines issued and signed by the Director of Research and Strategic Partnership for the Board.

Credit Author Statement

Halleluyah Aworinde, Segun Adebayo, Olufemi Alabi, Adebamiji Ayandiji, Echentama Kizito, Oke Olaide and Adetoye Adeyemo participated in data collection and labelling of the image dataset collected from Bowen University Research and Commercial Poultry Farms.

Segun Adebayo, Halleluyah Aworinde, Akinwale Akinwunmi, Aderonke Sakpere and Abel Oyebamiji segmented and annotated the audio signals, and took lead in writing and proof reading the manuscript.

All of the authors contributed to the manuscript and gave their approval to the final version after offering constructive criticism and helping to develop the research, analysis, and manuscript.

Data Availability

Poultry Vocalization Signal Dataset for Early Disease Detection (Original data) (Mendeley Data).

Declaration of Competing Interest

The authors declare that they are free of any financial or personal conflicts that might have had an effect on the research reported in this study.

Acknowledgments

References

[1] E.F. Briefer, Vocal expression of emotions in mammals: mechanisms of production and evidence, J. Zool. 288 (1) (2012) 1–20.
[2] N. Li, et al., Automated techniques for monitoring the behaviour and welfare of broilers and laying hens: towards the goal of precision livestock farming, Animal 14 (3) (2020) 617–625.
[3] E.S.M. Abdel-Kafy, et al., Sound analysis to predict the growth of turkeys, Animals 10 (5) (2020) 866.
[4] I. Fontana, et al., Sound analysis to model weight of broiler chickens, Poult. Sci. 96 (11) (2017) 3938–3943.
[5] K. Cuan, et al., Automatic Newcastle disease detection using sound technology and deep learning method, Comput. Electron. Agric. 194 (2022) 106740.
[6] A. Mahdavian, et al., Acoustic features of vocalization signal in poultry health monitoring, Appl. Acoust. 175 (2021) 107756.
[7] S. Adebayo, et al., Convolutional neural network-based crop disease detection model using transfer learning approach, Indones. J. Electr. Eng. Comput. Sci. 29 (1) (2023) 365–374.
[8] K. Cuan, et al., Gender determination of domestic chicks based on vocalization signals, Comput. Electron. Agric. 199 (2022) 107172.
[9] J. Lee, et al., Stress detection and classification of laying hens by sound analysis, Asian Australas. J. Anim. Sci. 28 (4) (2015) 592–598.
[10] X. Du, et al., Assessment of laying hens' thermal comfort using sound technology, Sensors 20 (2) (2020) 473.
[11] P. Chloupek, et al., Stress in broiler chickens due to acute noise exposure, Acta Vet. Brno 78 (1) (2009) 93–98.
[12] O. Halleluyah, et al., Poultry fecal imagery dataset for health status prediction: a case of South-West Nigeria, Data Br. 50 (2023) 109517, doi:10.1016/j.dib.2023.109517.
[13] J. Barros, R.I. Diego, On the use of the Hanning window for harmonic analysis in the standard framework, IEEE Trans. Power Deliv. 21 (1) (2005) 538–539.
[14] R.E. Kalman, A new approach to linear filtering and prediction problems, ASME. J. Basic Eng. 82 (1) (1960) 35–45.