An Acoustic Feature-Based Ensemble Learning Approach for Chicken Health Detection

Novita Rosyida
Department of Computer Science and Electronics, Universitas Gadjah Mada, Yogyakarta, Indonesia | Department of Creative and Digital Industry, Universitas Brawijaya, Malang, Indonesia
novitarosyida@mail.ugm.ac.id

Tri Kuntoro Priyambodo
Department of Computer Sciences and Electronics, Universitas Gadjah Mada, Yogyakarta, Indonesia
mastri@ugm.ac.id (corresponding author)

Afiahayati
Department of Computer Sciences and Electronics, Universitas Gadjah Mada, Yogyakarta, Indonesia
afia@ugm.ac.id

Zuprizal
Department of Animal Nutrition and Feed Science, Universitas Gadjah Mada, Yogyakarta, Indonesia
zuprizal@ugm.ac.id

Received: 24 December 2025 | Revised: 1 February 2026 | Accepted: 11 February 2026

 DOI: https://doi.org/10.48084/etasr.17149

ABSTRACT

Early disease detection in commercial poultry farms is critical for preventing outbreaks and minimizing economic losses. Conventional inspection is labor-intensive and frequently results in delayed diagnosis. This paper proposes an ensemble machine learning system for automated binary broiler health classification and evaluates the feasibility of non-invasive vocalization-based monitoring using acoustic analysis. Audio recordings were collected from 17–30-day-old broiler chickens in two closed-house commercial facilities: an academic research farm at the Faculty of Animal Science, Universitas Gadjah Mada (Yogyakarta), and a commercial farm in Blitar, Indonesia. Individual birds were recorded in isolated pens to eliminate background noise and ensure signal quality. A total of 22 acoustic features were extracted, comprising Mel-Frequency Cepstral Coefficients (13 features), time-domain features (5 features), and frequency-domain features (4 features). Three machine learning algorithms (SVM, Random Forest, and Logistic Regression) were evaluated across seven feature combinations using 5-fold cross-validation. Random Forest with MFCC features achieved the best individual performance (96.49% F1-score). An ensemble classifier with weighted soft voting was developed, integrating SVM (Time+MFCC), Logistic Regression (Time+Frequency+MFCC), and Random Forest (MFCC), with optimal weights determined through grid search, achieving 98.29% F1-score and 98.25% accuracy, outperforming the individual models. The high classification F1-score and accuracy demonstrate the feasibility of acoustic-based health monitoring for broiler chickens under controlled recording conditions to support early disease detection.

Keywords-acoustic monitoring; ensemble learning; poultry health; disease detection

I. INTRODUCTION

Maintaining the health and welfare of broiler chickens is critical to productivity and economic sustainability [1]. Respiratory diseases such as Infectious Bronchitis (IB), Infectious Laryngotracheitis (ILT), and Avian Influenza (AI) remain among the most severe threats, often spreading rapidly in intensive farms and causing high mortality and economic losses [2-4]. Early detection is essential for timely intervention and effective disease management. Recent advances in precision livestock farming have enabled the use of non-invasive sensors for automated health monitoring [5]. Among various sensing modalities, acoustic analysis has emerged as a particularly promising approach due to its ability to capture changes in poultry vocal behavior [6-8]. For instance, chickens affected by respiratory disorders may exhibit altered vocal patterns, including sneezing [3], coughing, and weak vocal intensity [9].

The application of machine learning techniques in agricultural automation has opened new possibilities for intelligent poultry farm management. Various classification algorithms have been explored for acoustic-based health monitoring, each offering distinct advantages in feature extraction and pattern recognition. Support Vector Machines (SVM) [10], Random Forest (RF) [11], and neural network architectures [4, 12] have demonstrated effectiveness in distinguishing healthy from diseased vocalizations through analysis of spectral and temporal acoustic features.

Ensemble learning is widely used to improve model reliability by combining several base classifiers into a single decision [13]. The main benefit typically appears when individual members are both reasonably accurate and sufficiently diverse, so that their mistakes are not concentrated on the same samples. Under these conditions, aggregation can reduce the effect of model-specific bias and variance, leading to better generalization than a single classifier [14]. In acoustic pattern recognition, similar gains have been reported when ensembles integrate evidence across complementary feature views or time scales, such as by combining multiple SVM classifiers trained on multi-time-scale aggregated speech features [15]. These observations motivate the use of ensemble learning for chicken sound-based health screening, where vocal signals can vary considerably across individuals, and recording conditions can challenge a single decision model.

In contrast, ensemble learning for broiler acoustic health classification remains comparatively underexplored, and ensemble configurations are often adopted without a systematic comparison of voting rules or weight selection strategies. Equal weight fusion or heuristic weighting is often adopted, and therefore, the practical advantage of optimized probability-based aggregation remains uncertain [16]. This paper addresses these gaps through a structured framework for non-invasive broiler health classification using vocalizations collected during the 17-30 day growth period. The main contributions are: (i) a unified evaluation of multiple acoustic feature sets and learning paradigms to clarify feature discriminability and algorithm feature interactions; (ii) development of a soft voting ensemble from complementary base learners with voting weights selected via grid search; and (iii) comparison with single classifiers, supported by confusion matrix based error analysis to quantify class specific mistakes relevant to health screening.

II. RESEARCH METHODOLOGY

This study followed a systematic experimental method comprising five main stages: (i) data collection from multiple commercial broiler facilities, (ii) audio preprocessing and feature extraction, (iii) individual classifier training and evaluation, (iv) ensemble construction with grid search optimization, and (v) performance evaluation. Figure 1 illustrates the complete research method, from data acquisition to the final classification results.

A. Data Collection

Audio recordings were collected from two geographically separated closed-house broiler farms to increase dataset variability and improve model generalization: (i) an academic research farm at the Faculty of Animal Science, Universitas Gadjah Mada (Yogyakarta, Indonesia), which operates under controlled management and intensive health monitoring; and (ii) a commercial production farm in Blitar, East Java, representing typical industry scale practices. Both sites used closed house ventilation with automated environmental control (temperature 28–32°C and relative humidity 60–70%).

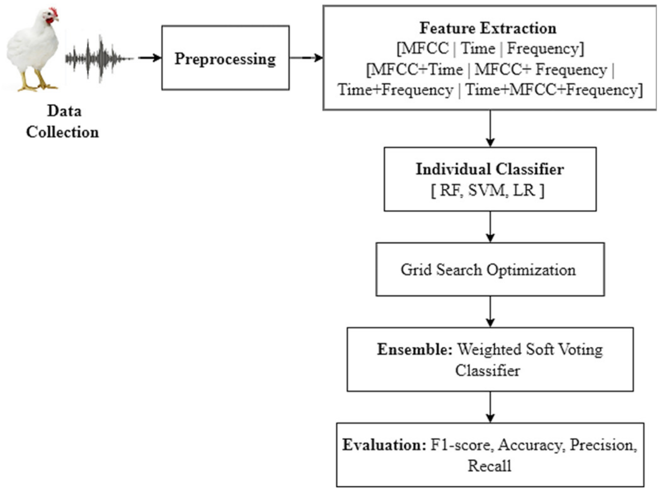

Fig. 1. Proposed method.

The study involved Cobb 500 broiler chickens aged 17 to 30 days, a critical early growth stage when maternal antibody protection declines and disease susceptibility increases. The health labels were assigned by experienced farm personnel using routine clinical observation protocols. Birds were labeled as healthy when they showed normal behavior, appetite, growth, feather condition, and activity, without visible clinical signs (e.g., respiratory distress, lethargy, abnormal posture, ocular/nasal discharge). Birds were labeled diseased when they exhibited clinical signs such as reduced feed intake, lethargy, respiratory symptoms (coughing, sneezing, labored breathing), abnormal vocalizations, or other behavioral changes. Classification was intentionally defined as binary (healthy vs. diseased) to support practical early-warning screening in commercial broiler production.

To ensure high signal quality and minimize acoustic interference, each bird was temporarily recorded in an isolated pen (100×100×60 cm). The audio data was captured in WAV format (44.1 kHz, 16-bit) at a fixed distance of 30–40 cm from the source. The recording sessions were conducted during the daytime (08:00–16:00), intentionally avoiding feeding periods and other management activities to minimize background noise. Furthermore, some birds were recorded over multiple days to capture natural within-bird variability in vocal behavior.

Following data collection, the raw audio streams were manually segmented into short clips of 2 to 4 s, each segment containing distinct chicken vocalizations. This process yielded a total of 1,140 segments, comprising 555 healthy chicken samples and 585 diseased chicken samples.

For the experimental phase, the dataset was rigorously partitioned into two sets: a training-validation set (80%, n = 912) and an independent hold-out test set (20%, n = 228). The independent test set, which was balanced with 114 healthy and 114 diseased segments, was strictly separated before training began to ensure the model's generalization capability and avoid data leakage. Finally, a 5-fold cross-validation was performed on the training-validation set to optimize the hyperparameters of the individual classifiers and determine the optimal weights for the ensemble model.

B. Audio Preprocessing

All recordings were resampled to a uniform sampling rate of 16 kHz. This rate was selected to satisfy the Nyquist-Shannon sampling theorem, since the primary frequency components of broiler vocalizations relevant for health screening (such as rales and sneezing) are generally below 8 kHz [17, 18]. The signal amplitude was then normalized to reduce the variability caused by the recording distance and the vocal intensity. Noise reduction was applied by STFT-based spectral subtraction using a noise profile reference, in which the estimated noise spectrum was subtracted from the noisy spectrum, and the signal was reconstructed back to the time domain [19]. Specifically, the signal was framed using a 25 ms analysis window with a 10 ms hop, employing a Hann window and *n_fft* = 512. The noise spectrum was estimated from low-energy frames using a percentile-based approach (*noise_percentile* = 10) and subtracted from the noisy magnitude spectrum with an over-subtraction factor $\alpha = 1.0$. To avoid musical noise artifacts and negative magnitudes, a spectral floor was applied using β = 0.02. The enhanced waveform was reconstructed via inverse STFT and used for subsequent feature extraction.

C. Feature Extraction

Comprehensive acoustic features were extracted using three complementary groups to capture spectral envelope, temporal/energy dynamics, and frequency distribution characteristics of broiler vocalizations. This multi-domain design is commonly adopted in animal/poultry sound recognition because health-related changes may appear as shifts in spectral shape (timbre), altered energy patterns, and reduced periodicity rather than a single cue [20]. Initially, 13 static MFCCs were computed to represent the short-term spectral envelope in a compact form, as MFCCs are a widely used and robust descriptor for audio pattern recognition and have been reported to be effective for poultry health monitoring [7].

To complement MFCCs, time-domain features were extracted, including Short Time Energy (STE) to represent intensity variations, Zero Crossing Rate (ZCR) to reflect voice changes, autocorrelation to capture quasi periodicity, Average Amplitude (AA) and Average Amplitude Difference (AAD) to summarize overall magnitude and short-lag variations, which are commonly used as efficient indicators of periodic structure and activity changes. Frequency-domain descriptors were included to directly quantify the spectral energy distribution: Spectral Centroid (SC) represents energy-weighted mean frequency, Spectral Entropy (SE) reflects concentration vs. spread/noise-like spectra, RMS frequency (energy-weighted frequency tendency), and mean spectrogram (time-averaged spectrogram magnitude capturing the long-term spectral profile). Table I provides a summary of the features extracted from each frame. For machine learning compatibility, frame-level features were aggregated into fixed-length vectors per recording to obtain consistent inputs across variable vocalization dynamics.

TABLE I. FEATURES SUMMARY

Feature type	Features	Dimension	Description
MFCC	13 static coefficient	13	Spectral envelope
Time domain	STE, ZCR, Autocorrelation, AA, AAD	5	Temporal characteristic
Frequency domain	SC, SE, RMS_freq, mean spectogram	4	Spectral distribution
Total (combined)	All feature	22	Comprehensive

D. Machine Learning and Ensemble Experimental Design

A two-stage evaluation strategy was applied to identify the most effective acoustic feature representation and classifier for broiler health screening and to assess whether combining complementary models improves robustness. In the first stage, an exhaustive comparison was conducted across three supervised learning algorithms and seven feature sets (MFCC, Time, Frequency, Time+MFCC, Time+Frequency, Frequency+MFCC, and Time+Frequency+MFCC) to ensure a fair and complete assessment of model and feature effects. The algorithms evaluated were Random Forest (RF), representing tree-based ensembles [21], Support Vector Machine (SVM), representing kernel-based classifiers [22], and Logistic Regression (LR), representing linear probabilistic models [23].

In the second stage, a soft voting ensemble was employed to investigate whether integrating models with diverse inductive biases and feature representations would enhance the robustness of the classification. The ensemble members were chosen using three criteria: a minimum performance threshold on cross-validation, algorithmic diversity across learning paradigms (tree-based, kernel-based, and linear), and feature diversity through complementary acoustic representations. The final class probabilities were derived from a weighted linear combination of the constituent classifiers, where the optimal weights were determined through a grid search subjected to a summation constraint of one. For comparison, the optimized weighting scheme was compared with equal weight voting, performance proportional weighting, and hard voting.

III. RESULTS AND DISCUSSION

A. Individual Model Performance Analysis

Table II summarizes the results of a systematic assessment of 21 algorithm and feature set combinations, obtained from three classifiers evaluated across seven feature representations. This comprehensive evaluation offers an overall view of performance, allowing informed identification of the most effective configurations, as well as meaningful insights into algorithm and feature interactions.

TABLE II. F1-SCORE PERFORMANCE FOR ALGORITHMS AND FEATURE SETS

Feature set	RF	SVM	LR	Best
MFCC	**96.49%**	93.86%	94.30%	**RF**
Time	92.54%	87.71%	88.16%	RF
Frequency	95.61%	88.14%	90.35%	RF
Time+MFCC	96.05%	**95.18%**	94.74%	RF
Frequency+MFCC	94.74%	**95.18%**	94.30%	RF
Time+Frequency	95.18%	87.70%	88.10%	RF
Time+Frequency+MFCC	94.74%	92.54%	**95.18%**	LR
Best	**96.49%**	95.18%	95.18%	96.49

Based on Table II, classification performance is strongly influenced by the selected feature representation, and the observed trends are consistent across the three evaluated classifiers. Overall, MFCC-based features provide the most reliable discrimination between healthy and diseased broiler vocalizations. The best result was achieved by RF with MFCC, reaching an F1-score of 96.49%, which indicates that spectral envelope information captured by MFCC is highly informative for broiler health screening. In contrast, the time domain feature set yields the lowest performance, particularly for SVM and LR, suggesting that temporal cues alone are insufficient when not supported by explicit spectral information.

From an algorithmic perspective, RF shows the strongest robustness, delivering the highest score for most feature sets, including both single-domain and combined representations. This behavior suggests that tree-based ensembles can effectively exploit heterogeneous acoustic descriptors and are less sensitive to the exact feature design. SVM performs best when MFCC is included, with the highest F1-score of 95.18% achieved for both Time+MFCC and Frequency+MFCC, highlighting that SVM benefits most from feature sets dominated by spectral content. LR reaches its peak performance on the full combination Time+Frequency+MFCC with an F1-score of 95.18%, indicating that additional complementary cues can improve a linear decision model when the feature representation is sufficiently rich. Combining feature domains did not consistently improve performance over MFCC alone. Although multi-domain sets increased dimensionality, they did not surpass the MFCC baseline, and in some cases produced lower F1-scores. This trend suggests that the additional descriptors may introduce redundancy or noise, and that the larger feature space can reduce generalization when the most informative spectral structure is already captured by MFCC.

B. *Ensemble Configuration and Member Selection*

The ensemble configuration was carried out after the single model experiments to obtain a classifier that is less dependent on one learning bias or one feature view. Candidate models were first filtered using the F1-score so that only consistently strong configurations were retained. From this shortlist, members were selected to balance two practical goals: maintaining high accuracy while combining models that tend to make different errors because they rely on different decision principles and feature representations.

TABLE III. ENSEMBLE MEMBER SELECTION SUMMARY

Model	Algorithm	Features	F1	Selection rationale
Member 1	SVM	Time+MFCC	95.18%	Kernel-based; temporal-spectral integration
Member 2	LR	Time+Freq+ MFCC	95.18%	Linear, comprehensive multi-domain
Member 3	RF	MFCC	96.49 %	Best overall; tree-based; spectral focus

The final ensemble consists of three members, as summarized in Table III. SVM with Time+MFCC was included to exploit nonlinear separations when temporal descriptors are added to the spectral envelope. LR with Time+Frequency+MFCC was chosen to add an interpretable linear probabilistic model that benefits from a richer multi-domain representation. RF with MFCC was selected because it provided the best overall performance and offers stable behavior on cepstral features. This combination satisfies performance requirements while preserving algorithmic and feature diversity, supporting the subsequent soft-voting stage and weight optimization.

C. *Ensemble Model Analysis*

Table IV presents the performance of four ensemble decision rules. The highest F1-score and accuracy are obtained with optimized weight soft-voting, achieving 98.29% F1-score and 98.25% accuracy. The weight vector [0.1429, 0.1429, 0.7143] was determined by an exhaustive grid search over feasible weight combinations subject to $\sum w_i = 1$, using the F1-score as the selection criterion. The resulting distribution assigns the largest contribution to the most reliable base learner while still preserving complementary information from the remaining members, yielding the best overall balance between precision and recall.

TABLE IV. VOTING STRATEGY PERFORMANCE COMPARISON

Strategy	Weight	F1	Acc.	Prec.	Rec.
Optimized Weight Soft-voting	[0.1429, 0.1429, 0.7143]	98.29%	98.25%	98.29%	98.29%
Equal Weight Soft-voting	[0.33, 0.33, 0.33]	96.14%	96.05%	96.55%	95.73%
Performance-proportional	[0.336, 0.332, 0.332]	96.14%	96.05%	96.55%	95.73%
Hard-voting	Majority rule	96.52%	96.49%	98.23%	94.87%

In comparison, the two baseline soft-voting schemes, namely equal weighting and performance proportional weighting, show nearly identical results with 96.05% accuracy and 96.14% F1-score, indicating that simple heuristics do not fully exploit the strengths of the ensemble members. Hard voting provides a modest improvement over these baselines in terms of F1-score and accuracy, but its recall decreases to 94.87%, which is undesirable for health screening because it increases the likelihood of missed diseased cases. Overall, the results demonstrate that probability aggregation with grid optimized weights offers a substantial advantage, improving F1-score by approximately 2.15% relative to the baseline voting strategies.

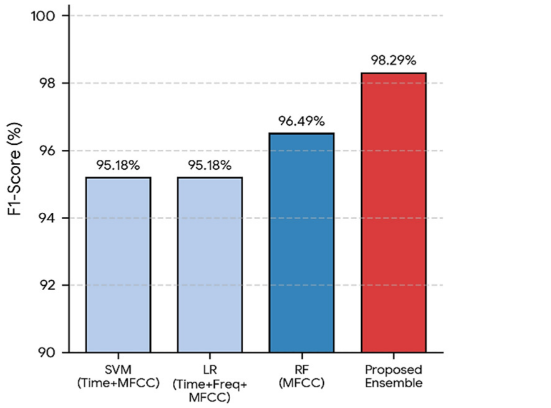

Fig. 2. F1-score performance comparison across different models and feature combinations.

Figure 2 illustrates the comparative results between the best-performing individual classifiers and the proposed ensemble model. Although RF with MFCC features provides a strong baseline with an F1-score of 96.49%, the ensemble approach using optimized weighted soft-voting notably improves the outcome to an F1-score of 98.29%. This graphical representation underscores how the fusion of multiple learning biases and complementary feature sets overcomes the limitations of standalone models, leading to more robust poultry health classification.

Validation of these results through confusion matrices (Figure 3) was conducted using an independent hold-out test set (n = 228). This dataset was kept strictly separate from the samples used during the 5-fold cross-validation training phase to provide an unbiased assessment. This visualization enables direct assessment of classification error patterns and demonstrates the ensemble's superior error correction capability, facilitating a direct evaluation of class-specific error distributions and demonstrating how probabilistic fusion enhances reliability beyond aggregate accuracy metrics. In the context of poultry health monitoring, false negatives, specifically misclassifying infected birds as healthy, constitute the most critical failure mode, as they risk delaying intervention and accelerating within-flock transmission.

Individual classifiers exhibit distinct performance trade-offs. While SVM and LR achieve comparable accuracy, both yield a higher incidence of false negatives, indicating compromised sensitivity toward diseased cases under the evaluated feature representations. In contrast, RF minimizes false negatives while maintaining a low false positive rate, suggesting superior separation between health states when utilizing the MFCC features. The proposed ensemble further refines the class-specific performance by suppressing both false negatives and false positives to a minimal count of two cases each. This translates into a recall of approximately 98.29% for the diseased class and a specificity of 98.20% for the healthy class, surpassing the metrics of standalone models. These findings suggest that the ensemble leverages complementary decision behaviors among its members, allowing misclassifications by single classifiers to be mitigated through probability level aggregation, thereby resulting in a more reliable screening outcome.

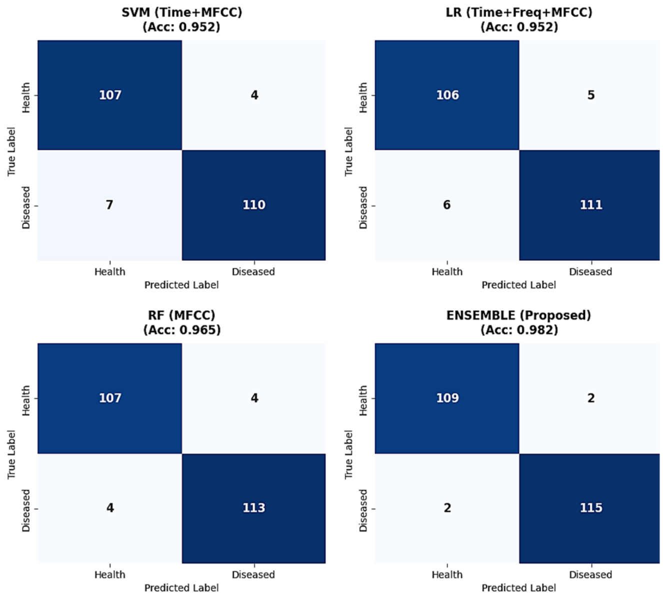

Fig. 3. Confusion matrices of the selected single classifiers and the proposed ensemble.

IV. CONCLUSION

This study introduces an automated framework for broiler health classification that integrates vocalization analysis with ensemble machine learning. The research provides a unified evaluation of multiple sets of acoustic features and learning paradigms, confirming that MFCC-based representations consistently offer the most robust separation between healthy and diseased vocalizations. Unlike the existing literature that predominantly utilizes individual classifiers or heuristic weighting schemes, the proposed ensemble approach achieves a superior 98.29% F1-score and 98.25% accuracy by employing grid-search optimized weighted soft-voting, effectively minimizing clinically critical false negatives to provide a more reliable automated poultry health screening tool. This method effectively halved the number of test errors from 8 to 4 compared to the best single-model baseline, significantly improving sensitivity for diseased cases. While the use of isolated recordings ensured high signal quality, future research will address the acoustic complexity of continuous barn environments and expand the scope to disease-specific identification for more precise farm management.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to Department Computer Science and Electronics, Faculty of Mathematics and Natural Sciences, and Faculty of Animal Science, Universitas Gadjah Mada, and the commercial poultry farm for their invaluable support and assistance during the data collection process. This research was supported by the Indonesian Education Scholarship (BPI) and the Indonesia Endowment Fund for Education (LPDP), under the administration of the Center for Higher Education Funding and Assessment (PPAPT), Ministry of Higher Education, Science, and Technology of the Republic of Indonesia, with decree number 01324/J5.2.3/BPI.06/9/2022.

DATA AVAILABILITY

The acoustic datasets generated and analyzed during this study are available from the corresponding author upon reasonable request. Requests will be fulfilled subject to institutional data sharing policies and ethical approval requirements.

REFERENCES

[1] A. Dhakal, S. Devkota, S. B. Jethara, R. K. Yadav, and P. Phuyal, "Assessment of Biosecurity in Poultry Farms in Chitwan, Nepal," *Veterinary Medicine and Science*, vol. 11, no. 2, Mar. 2025, Art. no. e70232, https://doi.org/10.1002/vms3.70232.

[2] A. Banakar, M. Sadeghi, and A. Shushtari, "An intelligent device for diagnosing avian diseases: Newcastle, infectious bronchitis, avian influenza," *Computers and Electronics in Agriculture*, vol. 127, pp. 744–753, Sept. 2016, https://doi.org/10.1016/j.compag.2016.08.006.

[3] L. Carpentier, E. Vranken, D. Berckmans, J. Paeshuyse, and T. Norton, "Development of sound-based poultry health monitoring tool for automated sneeze detection," *Computers and Electronics in Agriculture*, vol. 162, pp. 573–581, July 2019, https://doi.org/10.1016/j.compag.2019.05.013.

[4] 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.

[5] R. O. Ojo, A. O. Ajayi, H. A. Owolabi, L. O. Oyedele, and L. A. Akanbi, "Internet of Things and Machine Learning techniques in poultry health and welfare management: A systematic literature review," *Computers and Electronics in Agriculture*, vol. 200, Sept. 2022, Art. no. 107266, https://doi.org/10.1016/j.compag.2022.107266.

[6] J. Huang, T. Zhang, K. Cuan, and C. Fang, "An intelligent method for detecting poultry eating behaviour based on vocalization signals," *Computers and Electronics in Agriculture*, vol. 180, Jan. 2021, Art. no. 105884, https://doi.org/10.1016/j.compag.2020.105884.

[7] A. Mahdavian, S. Minaei, P. M. Marchetto, F. Almasganj, S. Rahimi, and C. Yang, "Acoustic features of vocalization signal in poultry health monitoring," *Applied Acoustics*, vol. 175, Apr. 2021, Art. no. 107756, https://doi.org/10.1016/j.apacoust.2020.107756.

[8] A. Mao *et al.*, "Automated identification of chicken distress vocalizations using deep learning models," *Journal of The Royal Society Interface*, vol. 19, no. 191, June 2022, Art. no. 20210921, https://doi.org/10.1098/rsif.2021.0921.

[9] Z. Sun, M. Zhang, J. Liu, J. Wang, Q. Wu, and G. Wang, "Research on white feather broiler health monitoring method based on sound detection and transfer learning," *Computers and Electronics in Agriculture*, vol. 214, Nov. 2023, Art. no. 108319, https://doi.org/10.1016/j.compag.2023.108319.

[10] 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.

[11] Z. Sun, W. Tao, M. Gao, M. Zhang, S. Song, and G. Wang, "Broiler health monitoring technology based on sound features and random forest," *Engineering Applications of Artificial Intelligence*, vol. 135, Sept. 2024, Art. no. 108849, https://doi.org/10.1016/j.engappai.2024.108849.

[12] K. Cuan, T. Zhang, Z. Li, J. Huang, Y. Ding, and C. Fang, "Automatic Newcastle disease detection using sound technology and deep learning method," *Computers and Electronics in Agriculture*, vol. 194, Mar. 2022, Art. no. 106740, https://doi.org/10.1016/j.compag.2022.106740.

[13] T. G. Dietterich, "Ensemble Methods in Machine Learning," in *Multiple Classifier Systems*, 2000, pp. 1–15, https://doi.org/10.1007/3-540-45014-9_1.

[14] A. Husin, "Designing Multiple Classifier Combinations: A Survey," *Journal of Theoretical and Applied Information Technology*, vol. 97, no. 20, pp. 2386–2405, Oct. 2019.

[15] A. Stefanowska and S. K. Zielinski, "Speech Emotion Recognition Using a Multi-Time-Scale Approach to Feature Aggregation and an Ensemble of SVM Classifiers," *Archives of Acoustics*, pp. 153–168, May 2024, https://doi.org/10.24425/aoa.2024.148784.

[16] H. Nawang, C. A. C. Yahaya, R. M. Sumeri, and K. A. Abdullah, "Weighted majority voting ensemble method for student performance classification," *International Journal of Advanced Technology and Engineering Exploration*, vol. 12, no. 129, Aug. 2025, https://doi.org/10.19101/IJATEE.2025.121220061.

[17] E. M. Hill, G. Koay, R. S. Heffner, and H. E. Heffner, "Audiogram of the chicken (Gallus gallus domesticus) from 2 Hz to 9 kHz," *Journal of Comparative Physiology A*, vol. 200, no. 10, pp. 863–870, Oct. 2014, https://doi.org/10.1007/s00359-014-0929-8.

[18] B. Krumm, G. M. Klump, C. Köppl, R. Beutelmann, and U. Langemann, "Chickens have excellent sound localization ability," *Journal of Experimental Biology*, vol. 225, no. 5, Mar. 2022, Art. no. jeb243601, https://doi.org/10.1242/jeb.243601.

[19] 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.

[20] W. Tao, G. Wang, Z. Sun, S. Xiao, Q. Wu, and M. Zhang, "Recognition Method for Broiler Sound Signals Based on Multi-Domain Sound Features and Classification Model," *Sensors*, vol. 22, no. 20, Oct. 2022, Art. no. 7935, https://doi.org/10.3390/s22207935.

[21] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001, https://doi.org/10.1023/A:1010933404324.

[22] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, Sept. 1995, https://doi.org/10.1007/BF00994018.

[23] C. Y. J. Peng, K. L. Lee, and G. M. Ingersoll, "An Introduction to Logistic Regression Analysis and Reporting," *The Journal of Educational Research*, vol. 96, no. 1, pp. 3–14, Sept. 2002, https://doi.org/10.1080/00220670209598786.