Poultry Health Classification using Deep Learning Based Analysis of Vocalization Signals Dataset

Meghavarshini M
School of Computing
SRM Institute of Science and Technology
Tiruchirappalli, India
meghxmohan@gmail.com

Mrishika D
School of Computing
SRM Institute of Science andTechnology
Tiruchirappalli, India
mrishikadhinakaran@gmail.com

***Abstract*— Rising protein demand has become a major driving force for poultry industry growth. However, the highest priority to ensure proper animal welfare and ensure a sustainable farm economy is actually keeping the health of the flock. Delayed diagnosis and considerable losses could be the consequences of traditional methods being based on time-consuming and erroneous manual observations. Yet, the innovative approach is one that makes use of artificial intelligence power. By registering and analysing vocalisations of poultry, the model can determine patterns associated with disease. This new system has several advantages: an early detection of diseases in their early forms, less utilization of labour costs for manual checking, increased accuracy in disease diagnosis, and substantial cost-effective savings. It promises to revolutionize poultry farming practices by offering secure flock health and sustainable food supply with the power of AI.**

Keywords—Poultry Health Monitoring, Audio Analysis, Deep Learning

I. INTRODUCTION

The poultry industry is crucial to global food security, but the sector faces recurrent problems due to disease outbreaks that impair productivity and animal welfare. The traditional approaches of monitoring health in poultry by manual observations and direct inspection are labor intensive, inconsistent, and unable in most cases to identify diseases at their onset. These limitations have spurred tremendous research into automatically scalable and non-invasive poultry health diagnostics powered by artificial intelligence (AI) [1][2].

AI-based solutions have great potential in conducting health monitoring in the poultry sector based on cutting-edge analysis of multimodal data, like audio and visual inputs. New studies have proven that deep learning (DL) models hold excellent potential in analyzing audio signals to distinguish anomalies in poultry health. Vocalizations can be regarded as a very important diagnostic tool since chickens indicate stress or illness by some changes in their vocalization patterns [4][5].For instance, the ML-based techniques can be utilized for audio recording analysis for the extraction of Mel-Frequency Cepstral Coefficients (MFCCs), and such analyses have been used for classification of poultry health state [1][8]. Besides the audio analysis, IoT-enabled systems with computer vision can be used in real-time poultry health monitoring.

They allow for the collection and processing of large amounts of datasets from agricultural settings to be able to accurately increase efficiency in health diagnostics [6]. Although other visual approaches, such as abnormal droppings, have also been added to DL models, audio and visual inputs were more robust to take on the challenges that life had in real farm conditions [4][5]. This work thus shines light on the possibility of using deep learning models in the analysis of poultry vocalizations, an emerging area for audio-based systems in the discovery of changes in disease status at early times. The contributions of this work fill gaps in existing approaches through scalable, accurate, and cost-effective solutions toward sustainable livestock management for advancing applications of AI in poultry farming.

This system captures the fundamental characteristics of vocal patterns by extracting Mel-Frequency Cepstral Coefficients from audio recordings followed by analysis on whether these characteristics match that of healthy or unhealthy avian subjects utilizing a CNN.

1.1 Contributions

Thus, in the wake of breakthrough discoveries in deep learning and audio processing, we have created a new poultry health monitoring system based on vocalization.

Audio recording from poultry is monitored to identify early health problems. The proposed CNN model takes advantage of 1D convolutional layers for recognizing poultry vocalizations, incorporating them as audio features in terms of MFCCs for efficient and accurate classification as healthy or unhealthy. This allows farmers to check the health status of flocks without physical examination. In this work, we have incorporated advanced feature extraction techniques with real-time prediction and have implemented the system with optimal usability.

We have used Python libraries such as Librosa for audio processing, and TensorFlow for model training and evaluation. Using techniques like early stopping, the system has been designed to achieve the best possible performance without overfitting. Further, the system has a web application available through smartphones, whereby farmers can upload audio recordings of poultry and access real-time health assessments. This ensures ease in usability and accessibility for the evaluation of bird flocks health from any location.

Designed for Practicality and Scalability: The system is designed to fit seamlessly into the world of poultry farming, from high-level operability to deployment at rural farms and industrial-scale operations. Automatic extraction of audio features to real-time health predictions through an interactive interface powered by Gradio, usability and scalability are thereby concentrated. Continuous, non-invasive monitoring along with actionable insights gives the poultry farming industry a step forward in managing diseases.

1.2 Organization of the paper

This document is structured in the following manner:

- Section I: Introduction
  Provides an overview of an emerging growth in demand for AI-based poultry health monitoring systems, with a focus on depicting the unavailability of the methods of disease detection traditionally, and key contributions of the project-the use of novel vocalization analysis for the purpose.

- Section II: Related Works
  The paper discusses the existing poultry health monitoring systems by current studies. It analyses sensor-based, image-based, and sound-based models. This section explains how the proposed system bridges the gap using audio features in disease detection.

- Section III: Methodology
  The following part describes the preparation of the dataset, including audio recordings of healthy and unhealthy poultry. It elaborates on the extraction of Mel-Frequency Cepstral Coefficients (MFCCs) as features and presents the architecture of the Convolutional Neural Network (CNN) model used for classification. Details of the model's training and the assessment methodology, are also discussed.

- Section IV: Results and Discussions
  Shares the experimental results including measures of training and validation loss as well as investigates the accuracy of the model.

- Section V: Conclusion and Future Directions
  Conclusion of key results and possible paths for further enhancement. There are activities involving database size increase for a variety of diverse poultry species, more audio feature inclusion, and enhancement of the web application's interface to be more accessible to the public.

- Section VI: References
  Contains a comprehensive compilation of all referenced materials, guaranteeing meticulous documentation of the sources utilized throughout the document.

This structured approach enhances comprehension and uniformity, allowing readers to effortlessly traverse the progression of the investigation from the initial problem declaration to the conclusions and possible future directions.

II. RELATED WORKS

Recent breakthroughs in artificial intelligence and deep learning have given remarkable transformation to poultry health monitoring systems that have brought the traditional problems of disease detection, productivity, and animal welfare to a considerable decrease. Audio analysis, IoT frameworks, and computer vision have lately emerged to be the latest focal points of research in this sector.

Audio analysis in recent times has emerged to be used in health monitoring in poultry as a means of diagnostic information obtained by the analysis of vocalizations in chickens and turkeys. Adebayo et al., (2023) applied the machine learning-based approach to process poultry voice signals using techniques that range from signal processing to classification of the health conditions in chicken [1]. Likewise, acoustic scene classification techniques, as reported by Abeßer (2020), highlight the potency of DL-based methods like CNNs in feature extraction from audio data into better monitoring systems in terms of health status [3]. Another research applied the diagnostic system based on deep learning to identify abnormal poultry faeces and vocalizations. There, the effectiveness in the application of various diagnostic modalities at early times for disease detection was indicated [4].

Capabilities of these systems are further enhanced by IoT-based frameworks in poultry health monitoring. An IoT-based health monitoring of chickens achieved and implemented in 2023 demonstrates how it is possible to efficiently process vast amounts of sensor network data towards timely anomaly detection [6]. Typically, multi-modal sources of data are involved: audio and video inputs enhance the robustness of the health assessment. For example, in work, Park et al. (2023) outlined how machine vision-based techniques could be combined with non-invasive diagnostics, thus demarcating the limitation of purely visual approaches in dealing with complex farm environments while further underlining the added benefits of having audio data [4][5].

Feature extraction techniques play a key role in audio-based diagnostics. Xia et al. (2019) present a broad survey of neural network-based deep learning techniques in acoustic event detection to highlight the gains in the accuracy of classification tasks via improvements in feature extraction and model architectures across diverse domains, including poultry health [7]. Equally, with audio signal processing and feature extraction trends as discussed in one paper published in 2020, it calls for refining methods such as MFCCs in order to capture health-related anomalies in vocalizations more efficiently [2]. Furthermore, it has explored unsupervised learning techniques and domain adaptation in order to enable real-time and scalable monitoring. This can be exemplified with the introduction of the unsupervised adversarial domain adaptation framework to account for audio data variability captured under multiple environmental conditions, thus improving acoustic scene classification models' generalization [9]. Such developments point to the need for robust and adaptive systems for trusting poultry health monitoring across farms' heterogeneous settings.

Although audio-based diagnostics are of great research interest, most potentially multimodal data combinations still have to be explored in more depth.

The combination of visual and audio inputs might lead to a more holistic evaluation of poultry health. Abnormal vocalizations often coincide with physical symptoms observable with computer vision techniques [4][5]. The adoption of such approaches would lead to a pathway toward having disease management and overall productivity enhancement solutions in poultry farming that are highly accurate, non-invasive, and cost-effective.

Altogether, these studies point out the promises of AI-based solutions to poultry health management. Audio signal integration, vision-based methods, IoT devices, and multimodal data all improve the accuracy and scalability of health monitoring systems but make them cost-effective and accessible. Pioneering researchers are opening pathways toward new and innovative solutions over challenges in poultry farming diseases and prevention, aiming to produce better productivity and animal welfare through the combination of different sensing technologies and various data modalities.

#	Study	Major Contribution	Limitation
1	Enhancing poultry health management through machine learning-based analysis of vocalization signals dataset (2023) [1]	Developed a machine learning-based system for analysing poultry vocalization signals to monitor health.	Focused primarily on the dataset and feature extraction, requiring further validation across different farm conditions.
2	Deep Learning-Based Poultry Health Diagnosis: Detecting Abnormal Feces and Analyzing Vocalizations (2024) [4]	Introduced a deep learning-based system for detecting abnormal feces and analyzing vocalizations in poultry health.	Focused on specific diseases, limiting generalization to broader health conditions.
3	Acoustic-Based Chicken Health Monitoring in Smart Poultry Farms (2023) [5]	Developed an IoT-based system integrating acoustic monitoring for real-time poultry health assessment.	Limited testing in real-world farm conditions, requiring further scalability analysis.
4	Exploring Deep Learning for Detection of Poultry Activities (2023) [6]	Investigated the use of deep learning to detect poultry activities for autonomous health monitoring.	Focused more on activity detection than direct health monitoring, limiting its immediate relevance to disease detection.

III. METHODOLOGY

Although much has been added in recent years to animal health monitoring systems, the models that currently describe classification of poultry health based on vocalizations are often inadequate in distinguishing healthy from unhealthy states. These systems have an underlying weakness where sensitivity in identifying relatively minor audio variability indicating distinct health conditions results in either false classification or the missed early signs of illness. Most of the solutions in literature rely on extensive computational resources to process and analyse the volume of audio files, which renders difficult to implement in standard hardware.

Our approach identifies several key innovations to solve these problems. Our approach uses CNNs, that employ Conv1D layers, to classify poultry health by extracting MFCC features from vocal recordings. This will enable more accurate and efficient classification even in resource-constrained low-computational conditions. We focus on developing a solution, while it is streamlined enough to be deployed easily in resource-constrained conditions, also providing high accuracy in the detection of poultry health states. Our goal is to develop an AI-driven system that can effectively classify poultry health from voice data, enabling farmers to get a low-cost and sensitive early warning signal about probable health issues.

3.1 Datasets

We used the Poultry Health Classification dataset downloaded from Mendeley as training data to train a deep learning classifier for classifying the health status of poultry from their vocalization signals. This dataset includes 346 audio files in total, categorized as healthy, unhealthy, and noise. Further specification of these categories is as follows: the category healthy has 139 audio files; the category unhealthy consists of 121 audio files, and the category noise comprises 86 files. All audio files are in .wav format.

A. Dataset Overview

- Sources and materials: The dataset was collected to analyse poultry vocalizations, which are good indicators of health. Captured on controlled environments from chickens that are suffering from both normal, healthy states and abnormal or unhealthy states, the sound source is made up from both healthy and unhealthy chickens. Moreover, it includes environmental noise to make it possible for the model to differentiate between the true poultry noises and the noise within the environments.

- Language: The audio files are in .wav format, a widely accepted format for audio data, ensuring high-quality sound recordings that are essential for accurate feature extraction and classification. The analysis is independent of language since it focuses on sound features, not spoken language.

- Categories: The dataset is divided into three primary categories. The healthy category includes audio recordings of chickens that are in good condition, with normal vocalizations free from any traces of distress or illness. The unhealthy category contains recordings taken from chickens displaying symptoms of illness, with possible stress or discomfort in their vocalizations. This includes noise category, which consists of environmental noise or background sounds that are not related to poultry vocalizations, thus allowing the model to differentiate sounds that are and are not relevant.

B. Data preparation

- Feature Extraction: To convert the raw audio data into a usable format for machine learning models, Mel-Frequency Cepstral Coefficients (MFCCs) were extracted from each audio file. MFCCs are commonly used features in speech and audio analysis as they capture the frequency characteristics of sound, which are essential for identifying patterns in vocalizations. For this dataset, 13 MFCCs were extracted from each audio file, providing a compact representation of the sound's spectral properties.

- Preprocessing: The audio files were processed to ensure consistency across the dataset. This included resampling all files to a sample rate of 22050 Hz, a standard for audio processing, and applying techniques such as padding to maintain consistent input lengths. The resulting MFCCs were then averaged across the time axis to generate fixed-length feature vectors for each audio sample.

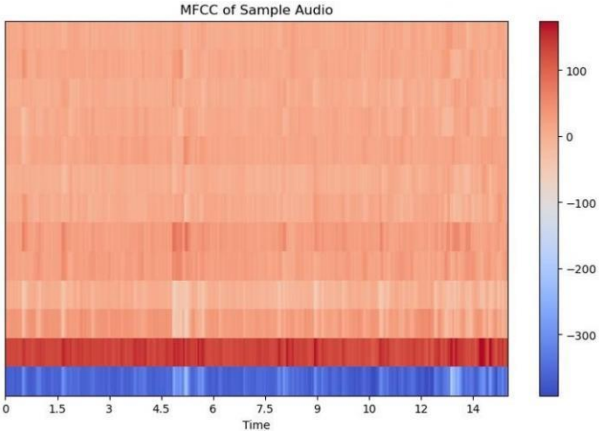

Fig 1: MFCC Visualization

C. Dataset Organization and Application

The dataset was specifically designed to support the development of machine learning models that classify poultry health based on their vocalizations. These three categories, namely healthy, unhealthy, and noise, primarily classify the input for training and evaluation of classification models.

Training and testing: This dataset was divided into training sets and testing sets to avoid overfitting, with 80% for training and 20% for the testing. In this way, a good balance is guaranteed because this model is tested on data which it hasn't seen in the overall training process, thus giving a well-rounded assessment of its strength. This resulted in the development of custom data distributions such that every class can represent itself well in both training and testing phases.

Model Training: The MFCCs extracted were further passed as the input of a deep learning-based model, which learnt to classify audio signals as representing a healthy, unhealthy, or noisy state. Hence, the model is expected to provide automated poultry health monitoring, which can be used in either farms or research environments to observe early signs of disease in poultry populations.

With careful organization and preparation of this dataset, it will thus constitute a good foundation for training accurate, efficient, and reliable models on poultry classification.

3.2 Proposed Architecture

Our front-end and back-end integrated poultry health monitoring system caters to farmers, providing a user-friendly, efficient, and scalable real-time solution for farm-based disease detection. The system makes use of accessibility using intuitive advanced machine learning techniques along with actionable insights.

Here are the key components of our system:

A. User Interface:

The Gradio dashboard serves as the primary interface with which users interact. It is simple, efficient and yet friendly to use by everyone, with technical knowledge not being a constraint to using the system.

- Farmers can upload audio recordings of poultry vocalizations or record audio on the web app directly, in widely accepted audio formats such as mp3 and .wav formats.
- Once the audio file is uploaded, the system processes the recording using the deep learning model, and gives the user the result.
- The result clearly classifies the health status of the poultry as either **"Healthy"** or **"Unhealthy"**, helping farmers get instant results about their avian health doubts.
- The dashboard interface is designed with a minimalist layout. It is designed keeping in mind the technical knowledge of the end user, with clear instructions at every step of the process.

B. Backend Elements

The backend, or framework, of the system contains crucial components that allow for model management and response generation.

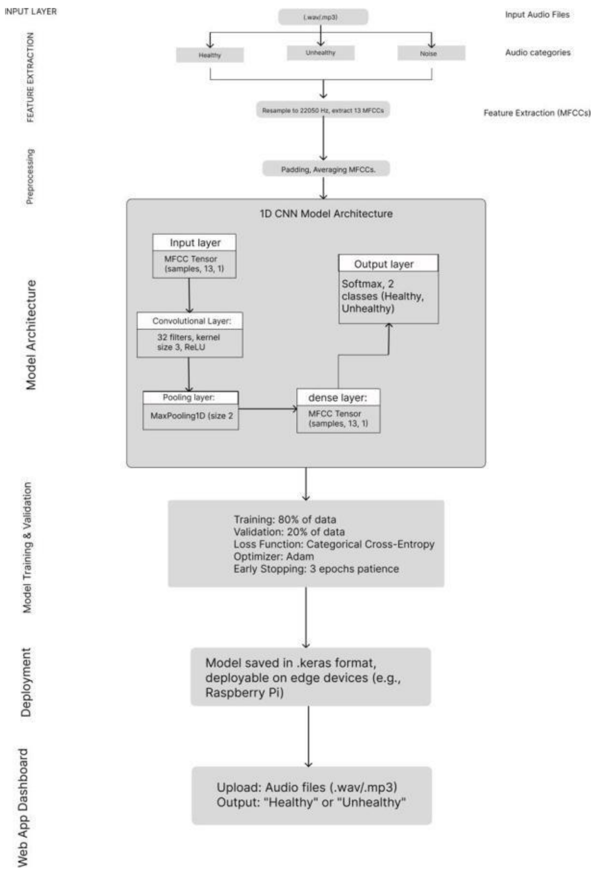

Fig. 1. Architecture diagram

The architecture opted for the experiment is a 1D Convolutional Neural Network, which is engineered specifically to be so for time-series data for feature extraction and classification from Mel-Frequency Cepstral Coefficients (MFCC) computed on audio files for broiler vocalization. This architecture is expected to work highly on accuracy in classification and scalable to being adaptable enough for real-time ability. Here is a step-by-step description of its major constituents:

- Input Layer: The model expects a three-dimensional input tensor in the shape (samples, num_MFCCs,1). Here, every sample refers to a processed audio recording. num_MFCCs is set to 13, referring to the number of MFCC features it has generated from each recording. The last dimension 1, that tells it is a single channel that is monophonic audio. This input configuration is suitable to represent the time patterns inherent in audio signals of poultry.

- The network contains a convolutional layer that has 32 filters of size 3. This is the most important layer of this network in order to pick-up local patterns within the MFCC features, such as abrupt changes in frequency, which could indicate health related occurrences.

- Activation Function: A ReLU activation function is used. It introduces non-linearity in the model so that it can learn abstract representations.

- Pooling Layer: A MaxPooling1D layer adds to min out the spatial dimensions and reduce computation overhead with a pool size of 2. This layer picks the highest features existing in tiny subregions of the feature map. That helps to retain whatever relevant information exists while discarding most of the redundant information.

- Flatten Layer: The pooled feature maps are forwarded to the Flattening layer. It unrolls the tensor of multidimensional levels into a vector in one dimension. This way, the data will be prepared for processing with the fully connected layers.

- Dense Layer: 128 neurons with ReLU activation. It's the high-capacity feature extractor that is designed to learn complex interactions amongst the MFCC features which correlate with health states.

- The final dense layer has a Softmax classifier that outputs the two health categories with probabilities: healthy and unhealthy. Because the softmax activation function ensures that the sum of output values equals 1, those output values can be interpreted as probabilities, and this causes the model to make robust and accurate predictions.

Model Training: The training phase refers to optimal adjustment of parameters for the desired accuracy with a deep CNN without overfitting. A combination of state-of-art techniques adapted here ensures reliable generalization to unseen data.

Loss Function: Categorical Cross-Entropy Loss is widely adopted for most of the multi-class classification problems. It computes the differences between predicted probabilities and true labels, thus driving the optimization process.

Optimizer: The Adam optimizer is used because of its adaptive learning rate, which automatically changes during training to increase the speed and stability of convergence. It perfectly fits with noisy gradients and sparse representations of data.

Early Stopping: A mechanism to avoid overfitting was incorporated in the applied use of early-stopping procedure. The training was stopped when validation loss didn't update for 3 consecutive epochs, thus preventing the model from just over-training on data and losing its ability to generalise.

3.3 Training Methodology

- Learning Rate: The default learning rate of the Adam optimizer, which varies dynamically during training for stable convergence, is 0.001.
- Batch Size: 32, chosen to balance gradient stability and computational efficiency.
- Epochs: 7, determined empirically to allow the model sufficient time to learn complex patterns in the data without overfitting.

3.4 Assessment Criteria

Confusion Matrix: Confusion matrix is one of the most important measures on how the model is classifying the audio files, across different categories. All that it does is show how often the model correctly makes labels and when it makes its mistakes. Ideally, most results should follow the diagonal, meaning that the model is making fewer errors and clear-cut distinctions between categories.

Validation Loss: Validation loss is monitored during training so that we know to what extent the model generalizes to unseen data. A general trend of decreasing validation loss in most cases would refer to learning generalizable patterns by the model. It can signify a problem like overfitting whenever the loss starts rising-that is, it then shows that the model is good on training examples but fails on new examples.

Accuracy: Accuracy is probably the simplest and most intuitive measure. It gives us the percentage of correct predictions that the model has made out of the total. The higher the accuracy, the more reliable and effective the model will be at choosing the right audio category.

AUC-ROC Curve: The AUC of the ROC curve is another important measure, which provides insight into how well a model can classify between different classes. The more that approaches 1, the better is the model's performance; that is, it is predicting correctly with a very high degree of accuracy for each class.

Class Distribution and Balancing: We also check that the model is trained equally across all classes so that there is no bias; if there is a considerable underrepresentation for some classes, it might affect the performance of the model. In this case, we could use data augmentation or balancing techniques to have fairer results and evaluate the model with a more accurate perspective.

IV. RESULTS AND DISCUSSIONS

The performance of the trained model was evaluated on several key metrics, including accuracy, loss, confusion matrix, receiver operating characteristic (ROC) curves, and class distribution analysis. The results provide insights into the efficacy of the model in handling the classification task, including its ability to generalize to unseen data and distinguish between classes.

4.1 Training Performance

Seven epochs of fine-tuning were used with the poultry audio health dataset. The main measures tracked during training were validation and training losses:

Epoch	Training Loss	Validation Loss
1	0.2704	0.2336
2	0.3194	0.2544
3	0.3406	0.2780
4	0.3054	0.2084
5	0.2776	0.2422
6	0.4347	0.2560
7	0.3707	0.2393

Fig 2: Training and Validation Losses

The loss values reveal an overall trend of convergence, with validation loss stabilizing after the fourth epoch. While fluctuations in training loss were observed in later epochs (e.g., Epoch 6 showed an increase to **0.4347**), the validation loss remained relatively stable. This suggests that the model maintained a strong generalization ability, despite minor instability in training.

The small rise in training loss in the last epoch points to possible dataset heterogeneity. This did not, however, compromise the validation performance, suggesting strong generalizing powers.

4.2 Model Performance

The accuracy of the derived model on the training set was 91.09% and had a corresponding loss of 0.2529. Test set accuracy dropped marginally to 90.0% while showing robust generalization capability with only minor performance drops; this suggests that the model actually captured the underlying patterns within the training data optimally, without causing significant overfitting.

4.3 Receiver Operating Characteristic (ROC) analysis

Here, for each class, the ROC curves are plotted as seen in the visualization. The AUC values for these three classes were: Class 0: 0.97, Class 1: 0.97, Class 2: 1.00. The AUC scores indicate excellent discrimination for all classes and perfect classification for Class 2. Nearly perfect AUC metrics for Classes 0 and 1 make it evident that the model is capable of distinguishing between positive and negative instances with high confidence. High AUC values are highly relevant in all situations where low false positives and low false negatives are critical.

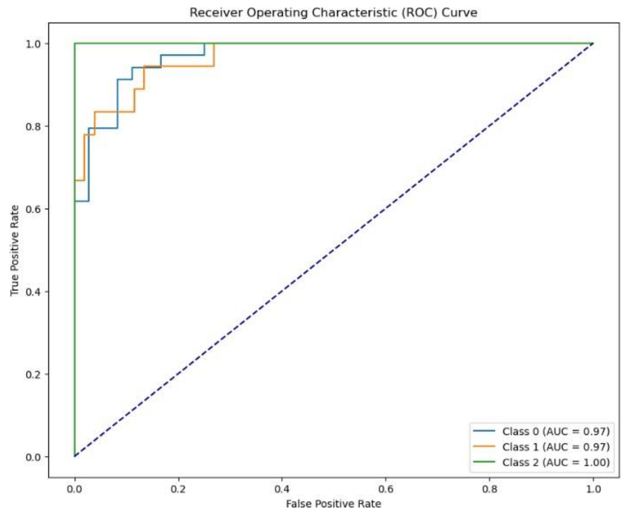

Fig 3: ROC Curve

4.4 Confusion Matrix Analysis:

A confusion matrix is a comprehensive description of the classification process by including true positives as well as false-positive, true-negative and false-negative predictions: Class 0 produced 33 actual positives and only 1 false positive. There were 12 true positives in class 1, but there were 6 misclassifications. Class 2 was entirely accurate having produced 18 true positives without any kind of misclassifications. The majority of errors were of Class 1, with a few taken as Class 0 wrongly. This is to be assigned to overlapping features of Classes 0 and 1, which implies that the features need to be engineered or augmented to enhance mutual separability.

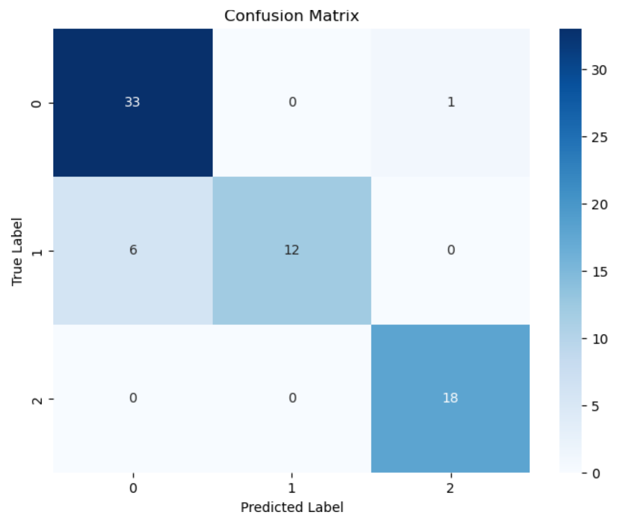

Fig 4: Confusion Matrix

4.5 Class Distribution

In class distribution analysis, it was found that all samples were equally distributed among classes. It is a requirement for equilibrium in class distribution to avoid bias in the model to a particular class. Imbalanced datasets mostly result in biased estimates and poor generalization performance due to the samples of one or more classes being significantly larger in number than others. Balanced datasets may have contributed to the raised performance metrics.

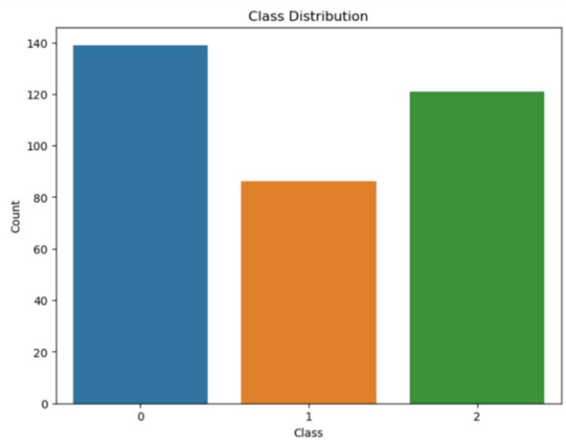

Fig 5: Class Distribution

The results reflect the robustness of the model in addressing the problem with the task of classification. High accuracy and AUC values indicate that the model could understand the nuances that exist within the dataset, whereas the confusion matrix points out some specific points to be addressed. Mistakes in Class 1 indicate that the model had difficulty not only in distinguishing various examples belonging to Class 0 from Class 1 but also it is probable that that occurs because of similarities in features in the dataset. Possible Directions: Feature Engineering: Introduce new features or modify the existing ones; this might improve separability between classes. This would help generate synthetic differences, mainly for those misclassified samples, hence better generalizing ability by the model.. Very good test accuracy scores of up to 90% accompany near-perfect AUC scores on all classes, thus making the model a very good candidate for practical implementation.

V. CONCLUSION AND FUTURE WORKS

5.1 Conclusion

It presents a revolutionary poultry health monitoring approach with deep learning-based analysis of vocalization signals. Traditional methods, which rely on inspections by humans, are tedious, prone to error, and could not diagnose the disease early. Therefore, the proposed system overcomes any limitations based on its 1D Convolutional Neural Network (CNN) and Mel-Frequency Cepstral Coefficients (MFCCs) regarding the proper classification of poultry health states.

This shows good performance, achieving 91 percent training accuracy and 90 percent test accuracy for the overall framework; thus, it can be said to pick up on health anomalies in real-world applications. Based on vocalization patterns, this system lends itself to a non-invasive efficient, and scalable early disease detection solution. It reduces the need for physical inspections as both costs of labour and stress on animals are decreased. Its light design ensures it can be deployed on standard hardware, thus being applicable to farms of all sizes, even in those resource-constrained environments.

The system is based on a convenient interface that is user-friendly, and it has the power of Gradio to it. It will facilitate uploading of audio recordings and allow getting instant health assessments, thus reduced effort. Through the explicit design of the dashboard, it ensures non-technical people can find their way around for actionable insight taken promptly to have timely interventions. The system is therefore very practical in modern poultry farming with integration into real-time monitoring and ease of use.

While direct implications exist for the farmers in terms of what this study brings to the farmers directly, it holds broader implication for food security and sustainability. Early detection of poultry diseases reduces the chances of an outbreak and protects productivity, which represents a stable food supply. It is an innovative audio analysis use for diagnostic purposes, standing as pioneering over traditional sensor- and vision-based methods and broadening the scope of precision agriculture.

This paper discusses the potential for revolutionizing poultry health monitoring with the aid of artificial intelligence. This work presents an easily scalable solution and an approach at an economical cost that coincides with animal welfare principles nowadays. The horizon for AI-based innovation in agriculture will steer systems like this toward much more sustainable and resilient food production. Future directions will include increasing datasets, especially through multimodal data sources, and improving system scalability so as to increase its impact on the poultry industry and further beyond.

5.2 Future Works

The proposed poultry health monitoring system is quite effective and innovative with a great potential for enhancement and extension. Future work should be focused on assessing the current limitations and extendibility of the system toward wider application and higher accuracy. Specifically, this enhancement is in the expansion of a dataset to include vastly more diversified variety of different audio recordings which will reflect a larger spectrum of poultry species, types of vocalizations, and conditions of environmental variations. This will make it strong and generalizable, meaning it works effectively in any farming-related setting. Feature engineering techniques can be advanced for further improvement in classification performance. Techniques such as spectral contrast and chroma features along with MFCCs may combine to capture subtle acoustic variations that are indicative of anomalies in health.

Another promising way forward is that this multimodal data from video feeds and sensor data from the Internet of Things must be integrated to give a more robust health monitoring system. For example, audio analysis could be combined with visual data or environmental metrics, such as temperature and humidity, that may give a deeper insight into those factors affecting poultry health. In addition, real-time alerts can be placed in place so interventions by farmers will be done in real time through SMS or IoT devices.

It would be essential to optimize the system for deployment on edge devices such as Raspberry Pi or its equivalent low-cost hardware with scalability and accessibility so that the system could function well in rural areas that have poor internet connections, thereby making it more usable in underserved farming communities.

Finally, long-term monitoring of the system may be instituted to verify that the system is adequate enough for a disease breakout prevention and improvements on animal welfare over time. This model will likely be used in a wide variety of species of livestock to prove its versatility and usefulness as a flexible general-purpose solution for monitoring livestock health, with these advantages further improving its potency role within sustainable agriculture, improved food security, and enhanced animal welfare at all levels around the world.

It is equally important for improvement in the user interface in uptake-this would involve, for example, extended multilingual support and introduction of visual analytics such as trend graphs and health status reports to help farmers intuitively monitor the health of their flock over time.

REFERENCES

[1] S. Adebayo, H. O. Aworinde, A. O. Akinwunmi, O. M. Alabi, A. Ayandiji, A. B. Sakpere, A. Adeyemo, A. K. Oyebamiji, O. Olaide, and E. Kizito, "Enhancing poultry health management through machine learning-based analysis of vocalization signals dataset," Data in Brief, vol. 50, p. 109528, 2023. doi: 10.1016/j.dib.2023.109528.

[2] "Trends in audio signal feature extraction methods," Applied Acoustics, vol. 158, p. 107020, 2020. doi: 10.1016/j.apacoust.2019.107020.

[3] J. Abeßer, "A review of deep learning-based methods for acoustic scene classification," Applied Sciences, vol. 10, no. 6, p. 2020, 2020. doi: 10.3390/app10062020.

[4] "Deep learning-based poultry health diagnosis: Detecting abnormal feces and analyzing vocalizations," IEEE Conference Publication, Apr. 17, 2024. Available: https://ieeexplore.ieee.org/document/10547723.

[5] "Acoustic-based chicken health monitoring in smart poultry farms," IEEE Conference Publication, Dec. 18, 2023. Available: https://ieeexplore.ieee.org/document/10467093.

[6] "Exploring deep learning for detection of poultry activities—Towards an autonomous health and welfare monitoring in poultry farms," IEEE Conference Publication, Jan. 3, 2023. Available: https://ieeexplore.ieee.org/document/10035656.

[7] X. Xia, R. Togneri, F. Sohel, Y. Zhao, and D. Huang, "A survey: Neural network-based deep learning for acoustic event detection," Circuits, Systems, and Signal Processing, vol. 38, no. 8, pp. 3433–3453, 2019. doi: 10.1007/s00034-019-01094-1.

[8] "Acoustic scene classification using deep audio feature and BLSTM network," IEEE Conference Publication, Jul. 1, 2018. Available: https://ieeexplore.ieee.org/document/8455765.

[9] "Unsupervised adversarial domain adaptation based on the Wasserstein distance for acoustic scene classification," IEEE Conference Publication, Oct. 1, 2019. Available: https://ieeexplore.ieee.org/abstract/document/8937231.