

# Multimodal Neural Network for Detecting and Classifying Deviations in Poultry Behavior

Garik Markarian  
R&D Department  
RiniSoft Ltd.  
Sliven, Bulgaria  
gm@rinisoft.com

Georgi Kolev  
R&D Department  
RiniSoft Ltd.  
Sliven, Bulgaria  
gk@rinisoft.com

Denis Kolev  
R&D Department  
RiniSoft Ltd.  
Sliven, Bulgaria  
dk@rinisoft.com

Natalia Polushkina  
R&D Department  
RiniSoft Ltd.  
Sliven, Bulgaria  
np@rinisoft.com

**Abstract** – This paper details the development and training of a multimodal neural network designed for the analysis of optical and acoustic data streams. Solution described here is one of the user products developed as a result of NESTLER project, a Horizon 2020 project which proposes the implementation of an environmentally sustainable, integrated technological and information solutions for agriculture. The primary outcome of this research is the establishment of the Poultry Quality of Life Index. This index enables the use of advanced mathematical diagnostics to evaluate and forecast the health status of poultry flocks. Poultry Health Monitoring System, developed based on the Poultry Quality of Life Index, enables timely interventions to minimize damage from the negative impact of external factors. By leveraging the PQL Index, the system can accurately evaluate the health status of the flock and promptly identify signs of distress or potential threats. By continuously monitoring the flock's health and environmental conditions, the system can detect early signs of distress or potential threats. This proactive approach allows for immediate corrective actions, thereby reducing the impact of negative influences on the flock's well-being and overall productivity.

**Keywords** – Multimodal neural network, optical and audio flows, digital poultry farm

## I. INTRODUCTION

Among various contemporary technical concepts, Digital Agriculture (DA) stands out as one of the most frequently utilized. DA leverages advanced technologies to enhance food production processes, increase efficiency and sustainability, and integrate information technologies and techniques into agricultural practices.

The main formalizations of Digital Agriculture [1] can be categorized as follows:

- *Precision Farming*: This approach relies on data collected through various sensors. For instance, information on soil humidity, temperature, and other environmental factors is used to optimize irrigation, fertilizer application, and numerous other aspects of agribusiness.
- *Robotic Agriculture*: This domain focuses on automating a wide range of agricultural tasks, from planting and harvesting to weed control and crop monitoring.
- *Artificial Intelligence (AI)*: AI in agriculture utilizes big data to provide actionable insights and recommendations to producers. By analyzing vast amounts of data, AI systems can suggest optimal times for planting, irrigation, and harvesting, and assess the health status of livestock populations. AI-

driven tools improve decision-making and enhance overall farm management.

- *Drones*: These are employed to monitor crops, identify diseases and pests, and map farms. Drones gather crucial information on plant and soil health, enabling precise and timely agricultural interventions.
- *Mobile Apps and Platforms*: These tools provide farmers with valuable information on farming operations, weather forecasts, market trends, and other essential resources. They facilitate informed decision-making and efficient farm management.

One significant area within Digital Agriculture is the digitalization of the poultry industry. The primary objective of digitalization in this sector is to enhance the efficiency of poultry farms. This includes increasing biomass production and egg yield while simultaneously reducing the costs associated with bird maintenance and minimizing the use of pharmaceuticals [2].

Improving the quality of life of poultry in farms is one approach to achieving these goals. Due to the nature of poultry production, a flock of birds housed in a single room is considered as a collective unit. Parameters of individual birds within this population are used to develop a generalized model representing the entire flock.

The assessment of poultry quality of life (health) is based on the detection and classification of deviations in flock behavior. A generalized Poultry Quality of Life Index (PQLI) is introduced to evaluate the health status of the flock. This index provides a comprehensive measure of the flock's well-being, enabling targeted interventions to enhance overall productivity and health.

The Poultry Quality of Life Index (PQLI) is a complex metric determined based on the physiological parameters of the monitored flock. Given the economic impracticality of individually monitoring each flock member, non-invasive methods of data collection, such as video surveillance cameras with integrated microphones, are utilized to monitor the physiological parameters of the birds within the flock.

To accurately determine the PQLI, the system relies on input data derived from optical streams and the acoustic environment at the flock site. This approach enables continuous and comprehensive monitoring of the flock's condition without the need for direct physical interaction with individual birds. By analyzing visual and auditory data, the system can detect and classify behavioural deviations, which are critical for assessing the overall health and well-being of the flock.

## II. STATEMENT OF THE PROBLEM

To utilize this information effectively, the Poultry Health Monitoring System (PHMS) was developed, with its core being a Multi-Modal Neural Network (2M2N). The Poultry Health Monitoring System Functional Diagram is shown in Figure 1. The primary output of the PHMS is the Poultry Quality of Life Index (PQLI), which allows farmers to evaluate the effectiveness of their rearing strategies, the quality of the living conditions, the adequacy of water supply, and the level of external biological threats to the flock.

In the Multi-Modal Neural Network (2M2N), acoustic and optical streams form the foundation for analyzing the behavior of domestic birds. The information obtained from these heterogeneous sources complements and validates each other, enhancing the overall reliability and significance of the data. Through synergetic information processing, the algorithm calculates a measure of proximity of the current state of the population to previously verified regions within the state space. This integrative approach allows for a comprehensive and accurate assessment of the flock's health and behavioral patterns, facilitating timely and informed interventions.

## III. PROPOSED SOLUTION

The PHMS is designed to assess poultry flock health through advanced analysis of multimedia data, focusing on early disease detection via physiological and behavioral characteristics. Monitoring devices, such as microphones and cameras, are utilized to capture vocalizations, activity, and posture.

The system architecture comprises multiple layers:

- *Data Collection Layer*: Includes IoT devices such as cameras and microphones.
- *Data Analysis Layer*: Encompasses preprocessing and analysis modules (Behavior Analyzer, Appearance Analyzer, Vocalization Analyzer) using Convolutional Neural Networks (CNNs) for image recognition and classification, Deep SORT for tracking, and specialized algorithms for vocalization analysis.
- *Decision-Making Layer (Poultry Health Assessor Module)*: Applies advanced analytics to derive actionable insights from the collected metrics and parameters.

For the chicken health assessment, the following technologies were used:

### A. Behavior Analysis Strategy

Addressing the challenges inherent in tracking individual chickens within a large flock in a confined area, we have revised our behavior analysis strategy. Instead of focusing on individual tracking, our approach now centers on analyzing the average behavior of the flock per video frame. This includes observing patterns in eating, sleeping, and resting [3].

To achieve this, we utilize models such as YOLO (You Only Look Once) [4] to detect chickens within the video frame. Once detected, these detections are cropped, and a secondary YOLO model is applied to identify specific parts of the chickens and construct skeletal representations. This dual-stage process enables a detailed assessment of the actions of each cropped chicken. For the current implementation, an off-

the-shelf YOLO v7 model was employed. The architecture of the model remained unchanged.

The training was conducted using a combination of the COCO dataset and our custom-labeled dataset, which comprises several thousand images of PQf chickens. No modifications or enhancements were made to the training process itself.

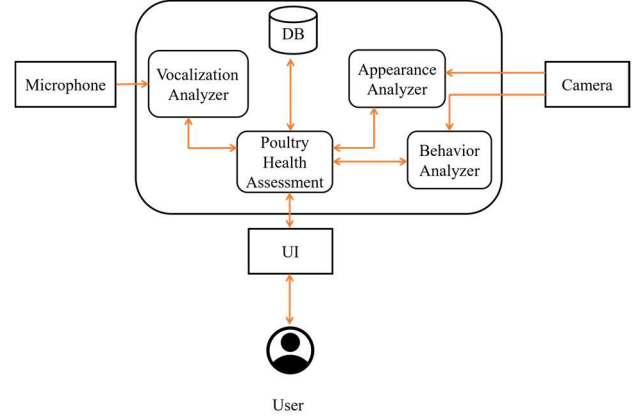


Fig. 1. Poultry Health Monitoring System Functional Diagram

### B. Training and Evaluation

The training of the YOLO v7 model spanned 80 epochs, with convergence monitored manually through TensorBoard, a utility that logs training metrics. An example of a training run on the combined COCO and chicken dataset is demonstrated in Figure 2.

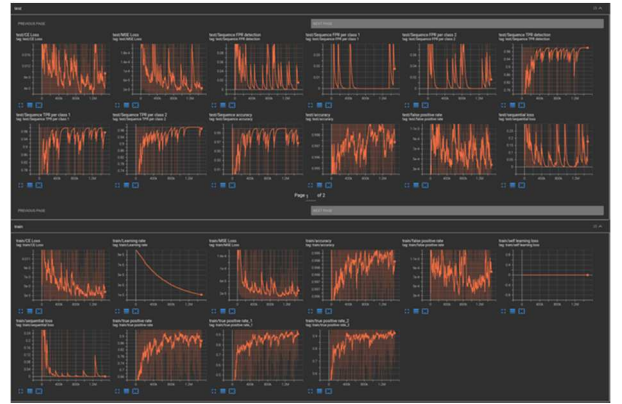


Fig. 2. YOLO v7 training tensorboard

The model's performance was evaluated on a test set comprising 197 annotated chicken images (totaling 2560 instances) and 5000 images from the COCO validation set, which were manually verified to contain no chickens. This was used to assess the false-positive rate. A chicken was considered detected with an Intersection over Union (IoU) threshold of 0.3, a slightly higher value chosen to ensure the reliability of the obtained metrics [5].

The evaluation yielded the following results:

- **True-Positive Rate:** 79.8%, corresponding to a 0.51% false-positive rate (Figure 3).
- **Average IoU for True Positives:**  $0.491 \pm 0.026$ .

The developed neural network was implemented in PyTorch and tested on an Nvidia Jetson Xavier board, a low-

power device (20-30W) featuring a GPU with 384 general-purpose cores and 48 tensor cores. Prior to performance testing, the trained YOLO model was exported to TensorFlow-1 protobuf graphs, allowing for 2-3 times faster inference compared to standard PyTorch or Keras formats. The inference on 1000 frames took  $59.3 \pm 1.5$  seconds.

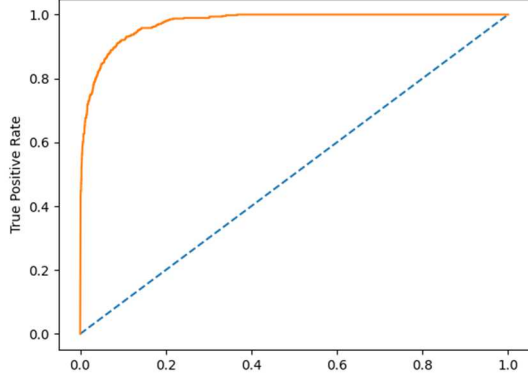


Fig. 3. ROC Curve

Future improvements to the detector could involve enhancing and expanding the visual training dataset and incorporating advanced training techniques such as MixUp, mosaic, and SimCLR, which are widely used to improve neural network performance.

#### IV. OPTICAL FLOW ANALYSIS FOR FLOCK MOBILITY

To detect anomalies in overall flock mobility, we have employed statistical optical flow analysis. Historical data and timestamps are recorded to correlate mobility patterns with specific times of day and year. This analysis currently does not utilize neural networks; instead, it focuses on identifying statistical outliers and assessing flock health based on threshold exceedances.

The algorithm implementation combines the YOLO v7 detector, as described previously, with optical flow analysis. We utilize the Farneback optical flow method, chosen for its balance between computational speed and estimation accuracy. The YOLO detector filters out detections unrelated to chickens, ensuring that the optical flow analysis focuses solely on the intended subjects.



Fig. 4. Example of an image and estimated speed/position PDF using KDE

During the set-up phase, the algorithm identifies the primary activity areas of the birds by employing Kernel Density Estimators (KDE) over the evaluated optical flow. This process involves a training period where the KDE maps are generated. Figure 4 provides a visual representation of the

density and movement patterns of the flock within the camera frame.

From the figures, it is clear that most activities occur in open areas and near the feeder. This information is captured by the estimated probability density function (PDF), defined over time and space. The spatial domain is represented in camera pixels, assuming a fixed position. The temporal domain is presented as a 24-hour periodic space, encoded using sine and cosine variables. Consequently, the final PDF is defined for three coordinates:  $(x, y, t)$ , which are implicitly translated to  $(x, y, \sin(t \bmod 24), \cos(t \bmod 24))$ , assuming a 24-hour time representation. To mitigate the impact of small-scale variations, the spatial resolution is coarsened, estimating the PDF for every three hours and for every  $200 \times 200$  pixels.

For each of these coordinates, two probability functions are defined:

- *Probability of Bird Presence*: Indicates the likelihood of birds being present in a given pixel.
- *Expected Optical Flow*: Estimates the expected motion within each pixel.

The optical flow PDF is estimated using Parzen's window approach, balancing evaluation speed and function smoothness. The window width is determined using Silverman's rule, optimizing the density estimation process.

As part of the optical flow analysis, the system reports the "flock movement quantity," measured in meters per hour (m/hour). This metric represents the cumulative distance travelled per hour by all active members of the flock. The flock movement quantity shown in Figure 5 serves as an indicator of the activity level within the population and is directly related to the calculation of the Poultry Quality of Life Index (PQLI).

#### V. PROCESSING AUDIO STREAMS

##### A. Audio Analysis

To analyze audio signals, we propose converting audio data into an image format. Spectrometric analysis is employed to create a visual representation of sound, utilizing quantization frequency of 48 kHz and an amplitude range from 0 to 20 kHz.

As illustrated in Figure 6, the obtained audio spectrograms are fully suitable for classification, analogous to the processing of optical streams. This approach significantly reduces the computational resources required for data processing. When defining the information matching policy within the knowledge base management wizard, each matching rule can be applied for simultaneous comparison of two records across the entire dataset. Records with matching scores exceeding a specified threshold are grouped into clusters in the matching results. These matching results are not directly added to the knowledge base but are used to fine-tune the matching rules and decisive rules for detecting deviations in swarm behavior.

The ResNet18 neural network is utilized to classify spectrograms. The network was trained using the dataset referenced in [6]. The neural network assigns each spectrogram to one of the following classes: "normal state," "presence of problems," or "noise." The classification results, along with a proximity measure, are then passed to M2N for further analysis.

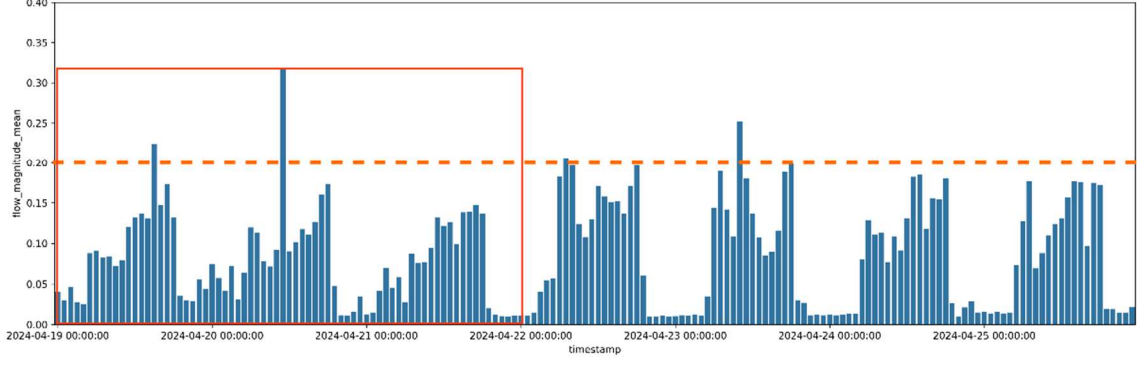


Fig. 5. Flock Movement Activity

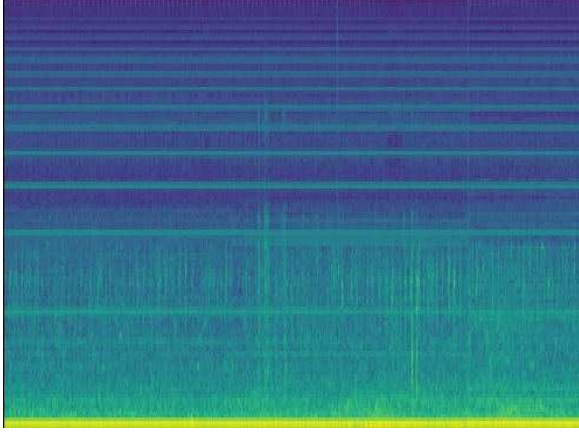


Fig. 6. Wheezing in the Respiratory System of Birds

The results of the 2M2N analysis using optical streams and audio spectrograms are represented as two-dimensional vectors, indicating the proximity measure to the verified swarm state on the same coordinate field. If both vectors fall within overlapping verified state areas, we can conclusively determine the state of the flock. For instance, if both wheezing sounds and reduced flock movement are detected, an unambiguous conclusion would be the presence of an infectious or viral disease among the birds [6].

The creation of matching policies and decisive rules is an interactive process wherein matching rules are iteratively modified based on the results of previous matches and statistical profiling data. This iterative refinement enhances the accuracy and effectiveness of the system in detecting and responding to deviations in flock behavior.

#### B. Determination of the State of Bird Population Based on 2M2N Work

The processes of detection, identification, and comparison are fundamental to human cognition and the recognition of various states of natural and anthropogenic objects. In the realm of mathematical diagnostics and discriminant analysis, the primary objective is to accurately determine the preliminary and true states of biological objects and to predict these states over specific time periods.

The data obtained from the operation of neural networks using audio and video streams is analyzed using a Decision Tree constructed in a direct manner.

In this context, the Poultry Quality of Life (PQL) Indexes serve as measures of the proximity of the current states of the tested objects to previously established and verified states. These proximity measures form the basis for constructing decisive rules, which are essential for state recognition and prediction.

The effectiveness of these proximity measures, and consequently the accuracy of the PQL Indexes, directly influences the correctness of state recognition and the reliability of future state predictions. By leveraging these indexes, we can establish a robust framework for assessing and forecasting the health and well-being of biological populations, such as poultry flocks [7].

The accuracy and utility of the constructed proximity measures and PQL Indexes determine the overall success of the system in recognizing and predicting the states of the biological objects being studied.

#### VI. FURTHER DEVELOPMENT OF THE PROPOSED SOLUTION

As part of the system's ongoing development, the next phase involves estimating physiological parameters of the flock, such as average weight gain, deviations from the average speed of movement considering external factors, and average feeding frequency. This will be based on synthesized metrics derived from optical flows and spectrograms.

#### VII. CONCLUSION

The developed PHMS utilizing a multimodal neural network offers a significant advancement in poultry health monitoring. By integrating optical and acoustic data streams, the system provides a comprehensive and continuous assessment of flock health, enabling timely and effective interventions. Future research will focus on further improving the accuracy of the PQLI and expanding the system's capabilities to include more diverse environmental and physiological factors.

#### VIII. ACKNOWLEDGEMENT

This research was supported by NESTLER project through European Commission Funding Grant Agreement No. 101060762 <https://nestler-project.eu/>.



## REFERENCES

- [1] Data Conceptual Model for Smart Poultry Farm Management System Nikolajs Bumanisa, Irina Arhipovaa, Liga Pauraa, Gatis Vitolsa, Liga Jankovskaba Latvia University of Life Sciences and Technologies, Liela street 2, Jelgava, LV, 3rd International Conference on Industry 4.0 and Smart Manufacturing 2022
- [2] Smart poultry house monitoring system using IoT R.Sasirekha1, Kaviya.R1, Saranya.G1Aezeden Mohamed, Utaeva Iroda E3S Web of Conferences 399, 04055 (2023) <https://doi.org/10.1051/e3sconf/202339904055>, ICONNECT-2023
- [3] A systematic review of machine learning techniques for cattle identification: Datasets, methods and future directions, Md Ekramul Hossain a e, Muhammad Ashad Kabir ab,e\*, Lihong Zheng a,e, Dave L. Swain, Shawn McGrath, Jonathan Medway, Artificial Intelligence in Agriculture 6 (2022) 138-155p
- [4] YOLOv7: A Powerful Object Detection Algorithm (2024 Guide). <https://viso.ai/deep-learning/yolov7-guide/> Sound Spectrum Analyzer, <https://academo.org/demos/spectrum-analyzer/>
- [5] Edge Computing and Artificial Intelligence for Real-time Poultry Monitoring Olivier Debauchea,b,, Sa'id Mahmoudia, Sidi Ahmed Mahmoudia, Pierre Mannebacka, J'er'ome Bindelleb,c, Fr'ed'eric Lebeaub. International Workshop on Artificial Intelligence & Internet of Things (A2IoT) August 9-12, 2020, Leuven, Belgium.
- [6] Poultry Vocalization Signal Dataset for Early Disease Detection, Published: 12 June 2023 Version 1, Contributors: Halleluyah Aworinde, Segun Adebayo, Akinwale Akinwunmi, Olufemi Alabi, Adebamiji Ayandiji, Olaide Oke, Abel Oyebamiji, Adetoye Adeyemo, Aderonke Sakpere, Kizito Echetam
- [7] Uzdin D.Z. On a new approach to the theory of state recognition. New methods of mathematical diagnostics. – M.: MAX Press, 2012