Open Journal of Physical Science (OJPS)
ISSN: 2734-2123
Article Details:
DOI: 10.52417/ojps.v6i1.762
Article Ref. No.: OJPS0601002-762
Volume: 6; Issue: 1, Pages: 01 - 13 (2025)
Accepted Date: 10th February, 2025
© 2025 Yakubu *et al.*

Open Journals Nigeria (OJN)
Open Access | Bi-annual | Peer-reviewed
www.openjournalsnigeria.org.ng
editorial@openjournalsnigeria.org.ng

RESEARCH ARTICLE **OJPS0601002-762**

UNMANNED AERIAL VEHICLE (UAV)-BASED COMPUTER VISION MODEL FOR REAL-TIME BIRDS DETECTION IN RICE FARM.

[*1]Yakubu, I., [1]Audu, M. M. & [2]Zayyanu, Y.

[1]*Computer Science Department, Yobe state University Damaturu, Nigeria*
[2]*Department of Computer Science, Bayero University Kano, Nigeria*

***Corresponding Author Email:** yakubu.ibrahim@ysu.edu.ng*

ABSTRACT

Rice farming in Nigeria suffers significant losses due to bird damage, necessitating advanced mitigation strategies. This study investigates the integration of computer vision with unmanned aerial vehicles (UAVs) to provide real-time bird detection and deterrence in rice fields. Given the varied agricultural conditions in Nigeria—including different farm sizes, vegetation density, and lighting conditions—the proposed system was designed for adaptability and robustness. Utilizing a dataset of 1,113 bird images captured via UAVs and ground cameras, a YOLOv8 model was trained with rigorous preprocessing and augmentation techniques. The model achieved a precision of 85%, recall of 70%, and mAP@50 of 80%, demonstrating strong detection capabilities. However, performance decreased in densely occluded environments, with mAP@50:95 stabilizing at 39%. Real-time testing confirmed the system's practical applicability and reliability under diverse environmental conditions. This solution represents a cost-effective, scalable approach to protecting rice fields, offering a significant leap forward in precision agriculture.

Keywords: *Unmanned Ariel vehicle, Computer vision, Deep learning, BIrds detection, Yolo model, Rice farm.*

INTRODUCTION

Rice farming is a vital agricultural activity in Nigeria, and plays a crucial role in ensuring national food security and strengthening the economy. (Mey & Demont, 2014). Crop losses due to bird infestations pose a major challenge to farmers in Nigeria, leading to substantial reductions in crop yields and economic returns (Roihan *et al.*, 2020). In rice fields, birds such as quelea birds, sparrows, and crows are notorious for causing extensive damage by feeding on seeds and mature grains (Agossou *et al.*, 2022). This issue results in considerable losses for Nigerian farmers who rely heavily on rice as a staple food and a crucial source of income (Okoli & Ezui, 2012). Traditional methods of bird control in Nigeria, including scarecrows, netting, and auditory deterrents, often prove to be ineffective or labor-intensive, providing only temporary relief. The need for a more reliable, efficient, and automated solution to protect rice fields from bird damage has never been more urgent (Riya *et al.*, 2020).

Recent advancements in technology offer promising solutions to this problem. Computer vision, a branch of artificial intelligence, allows machines to analyze and interpret visual information in a manner akin to human perception.. When combined with unmanned aerial vehicles (UAVs), more commonly referred to as drones., computer vision systems can offer real-time monitoring and intervention capabilities (Coluccia *et al.*, 2021). UAVs equipped with computer vision systems can patrol fields, detect bird activity, and initiate scaring mechanisms, providing continuous protection without the necessity for continuous human intervention (Marcoň *et al.*, 2021).

This study seeks to overcome these issues by developing and deploying a UAV-based computer vision model tailored specifically for the Nigerian agricultural landscape. The goal is to create a reliable, efficient, and cost-effective solution that can significantly reduce bird-related crop damage, thereby enhancing crop yields and supporting the livelihoods of Nigerian farmers.

To develop and implement an autonomous system using computer vision and UAV technology for real-time detection and scaring of birds in Nigerian rice fields, to reduce crop damage and increase yields. The following objectives have been set: 1. To collect and preprocess images of birds in the rice fields. 2. To develop advanced computer vision algorithms for real-time bird detection in rice fields. 3. To optimize and evaluate the effectiveness of the model in detecting birds in rice time from rice fields.

Linz *et al.* (2015) evaluated UAVs equipped with audio and visual deterrents for bird control in agricultural fields, showing significant reductions in bird activity. Wang and Wong (2018) combined machine learning and UAVs to protect vineyards from bird flocks, with a ground detection system prompting UAVs to track and deter birds. Their system was further improved in 2019 by enabling multiple UAVs to collaborate in deterring flocks. Guo *et al.* (2018) demonstrated the effectiveness of deep learning for pest detection, highlighting computer vision's potential for real-time field monitoring. Qiao *et al.* (2020) developed a deep learning-based model for bird species identification, further supporting the applicability of computer vision in agriculture. Z. Wang (2021) created a UAV-based system for bird detection in vineyards, showing high accuracy and effective deterrence.

Bhusal *et al.* (2022) automated a UAV-based bird deterrence system by integrating real-time bird movement detection, and UAV missions for intercepting and patrolling birds, and tested its efficiency in outdoor experiments. Marcoň *et al.* (2021) used a convolutional neural network (CNN) for detecting bird flocks and activating scaring devices only when needed, offering a solution to traditional deterrent habituation. Schiano *et al.* (2022) applied a CNN model with drones to autonomously detect and deter pigeons from rooftops. Bhusal *et al.* (2022) also highlighted the effectiveness of their system in vineyard trials, with successful mission completion rates of 92% for short tasks and 90% for longer ones. Acharya *et al.* (2024) introduced an AI framework for identifying BSLs in video footage using deep learning techniques like Faster RCNN, achieving a detection precision of 0.87 in varied weather conditions.

Despite these advances, gaps remain in adapting these systems to Nigerian rice fields. Key challenges include variable lighting, diverse bird species, and field conditions that may affect the reliability of existing computer vision models. Further research is needed to develop systems robust enough to handle the aforementioned unique agricultural conditions of Nigeria. This study aims to develop an autonomous system using computer vision and UAV technology for real-time detection and scaring of birds in Nigerian rice fields, to reduce crop damage and increase yields.

MATERIALS AND METHODS

Data Collection

S150 WiFi 4K HD Video Camera Drone was used to provide aerial views of the birds in various activities (feeding, flying) on the rice farms. Its high-resolution camera was crucial for capturing clear images, even in motion. Infinix X6528 (HOT40i) mobile phone was used for ground-level imaging, focusing on closer shots of birds. In total, 1113 images were captured across different rice farms, to ensure a diverse dataset that includes broad bird activities and perspectives, among which 543 were selected.

Data Preprocessing

Preprocessing was a critical phase to ensure that the dataset was ready for training the YOLOv8 model. The following preprocessing techniques were applied:

Images Resizing: All images were resized to 640×640 pixels using the OpenCV library in Python. This consistent image size reduces computational overhead during model training and ensures uniformity.

Data Augmentation: Various image augmentation techniques were applied to improve model generalization and simulate real-world variations:

- Horizontal Flipping: To account for symmetry in bird appearances.
- Random Rotations: Within ±15 degrees to simulate different orientations of birds.
- Brightness and Contrast Adjustments: To mimic varying lighting conditions.
- Cropping and Scaling: To handle situations with occlusions.
- Gaussian Noise: Added to simulate environmental distortions and camera artifacts. These augmentations were implemented using the Albumentations library, which is known for its efficiency.

Data Annotation: The images were annotated manually using Roboflow, a user-friendly tool. Bounding boxes were drawn around birds, labeling their positions whether they were flying or feeding. The annotations were saved in the YOLO format, which includes class, bounding box coordinates, and dimensions, all normalized to the image size.

Normalization: The pixel values were scaled to the range [0,1] by normalizing them through division by 255. This normalization step is important for faster convergence during model training.

Dataset Splitting: The dataset was partitioned into three sets to facilitate effective training, validation, and testing:

- Training (70%): Used for model training.
- Validation (20%): Used to evaluate model performance during training.
- Testing (10%): Reserved for final evaluation of the trained model.

Care was taken to maintain a balanced distribution of bird species and activities across these subsets.

Image Quality Enhancement: Images with poor quality or artifacts were identified and enhanced using Gaussian blurring to reduce noise. The resolution of low-quality images was also improved using OpenCV functions to ensure clarity and adequate quality for feature extraction.

Selection of YOLOv8 Model

For this research, We chose YOLOv8 for its high efficiency and accuracy in object detection tasks, balancing speed and accuracy, which makes it suitable for real-time applications (Yao *et al.*, 2024). YOLOv8 also allows custom training, making it adaptable to our bird detection task on rice farms. We specifically chose the YOLOv8s variant for its optimal balance between processing efficiency and detection accuracy, ideal for deployment on devices with small hardware resources.

Environment Setup: Training was conducted using Google Colab, a cloud platform that offers free GPU access. The environment setup included Google Colab Pro for extended runtime and access to the NVIDIA Tesla T4 GPU for accelerated training. The implementation was done using Python 3.10, with the Ultralytics library for YOLOv8 and PyTorch as the deep learning framework. CUDA was used to enable GPU-based training.

Model Training Process

We used the yolov8n.pt pre-trained model as a starting point for training on our custom bird detection dataset. All images were resized to 640x640 pixels to standardize input dimensions. The model was trained with a batch size of 8 over a hundred epochs and a predefined learning rate of 0.01. The SGD optimizer and YOLO loss function were used, and the dataset was divided into 70%, 20%, and 10% for training, testing, and validation, respectively. The model was set to predict based on an assurance threshold of 0.5 and an IoU threshold of 0.45 for Non-Maximum Suppression (NMS). The performance of our model was evaluated using metrics like precision, recall, and mAP@50, with pre-trained weights initialized before custom training.

Evaluation Metrics

To measure the performance of the training, it is necessary to use an extensive set of evaluation metrics. These metrics will help in understanding the model's accuracy, precision, recall, and overall effectiveness in various conditions. By using these evaluation metrics, the performance of the bird detection model can be thoroughly assessed and optimized, ensuring it meets the project's requirements for accuracy, efficiency, and robustness in real-world applications. The Key Evaluation Metrics to be used are:

Accuracy: Accuracy is the ratio of correctly predicted instances (both true positives and true negatives) to the total instances.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

Precision: The ratio of true positive detections to the total number of positive detections (true positives and false positives).

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

Recall (Sensitivity): The ratio of true positive detections to the total number of actual positives (true positives and false negatives).

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

F1-Score Is The harmonic mean of precision and recall, providing a balance between the two.

$$F1\ Score = \frac{Precision * Recall}{Precision+Recall} \tag{4}$$

Mean average Precision(mAP): measures the precision and recall performance of a model across multiple queries or classes.

$$\mathrm{mAP} = \frac{1}{n}\sum_{i=1}^{n} \mathrm{AP}_i \tag{5}$$

Where *n* is the number of classes and AP is the average precision of the classes.

RESULT AND DISCUSSION

We analyze the outcomes obtained both during and after the training of the model. This includes an evaluation of key performance metrics such as accuracy, precision, recall, F1-score, and loss trends. Visualizations, such as confusion matrices, precision-recall curves, and training-validation loss plots, are provided to illustrate the model's behavior and performance. Additionally, we performed real-time inference to validate the model's practical applicability and robustness in real-world scenarios. The results from these experiments are discussed in detail, highlighting the model's strengths, any observed limitations, and areas for potential improvement.

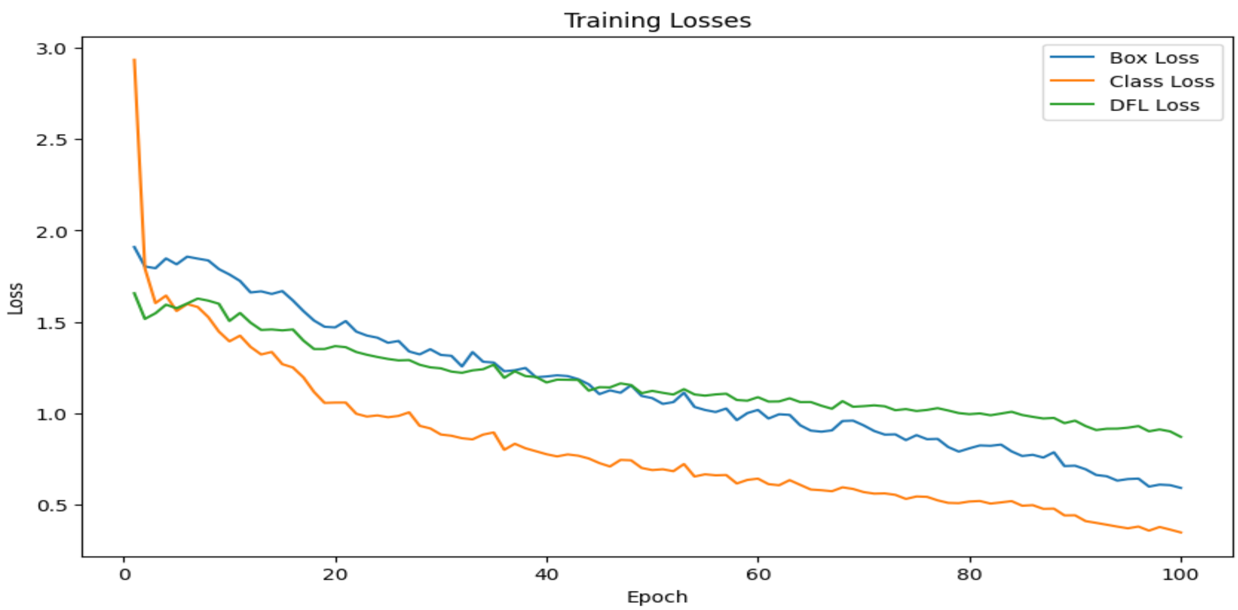

Figure 1: Training Losses

The above graph visualizes the training losses for the real-time bird detection model over 100 epochs, showing the following: Box Loss (Blue), Indicates inaccuracies in estimating the bounding boxes of identified objects. Class Loss (Orange), Evaluates the misclassification rate of detected objects into their appropriate categories. DFL Loss (Green): Refers to the distribution focal loss, which helps improve localization by adjusting the precision of bounding box predictions. All three losses decrease steadily, showing that the model is learning effectively. The losses plateau towards the later epochs, suggesting convergence and diminishing returns on further training.

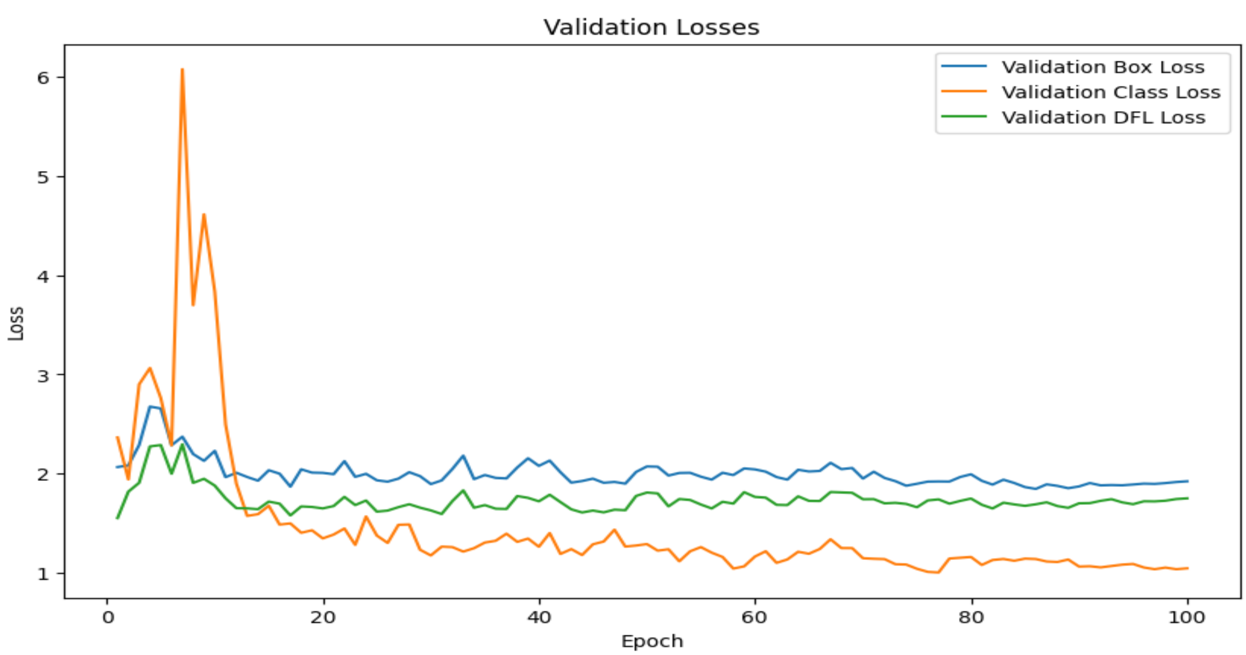

Figure 2: Validation Losses

The above graph illustrates the validation losses for the bird detection model over 100 epochs, including Validation Box Loss (Blue): This represents errors in predicting bounding boxes for validation data. Fluctuates slightly but remains steady after initial epochs, suggesting stability in localization predictions. Validation Class Loss (Orange), Indicates errors in classification for validation data. Validation DFL Loss (Green): Refers to the distribution focal loss during validation, helping improve the precision of bounding boxes. Validation losses exceed training losses, which is expected since the model is tested on previously unseen data. Stabilization in later epochs suggests the model is not overfitting. Minor fluctuations in validation losses indicate noise or challenging samples in the validation set.

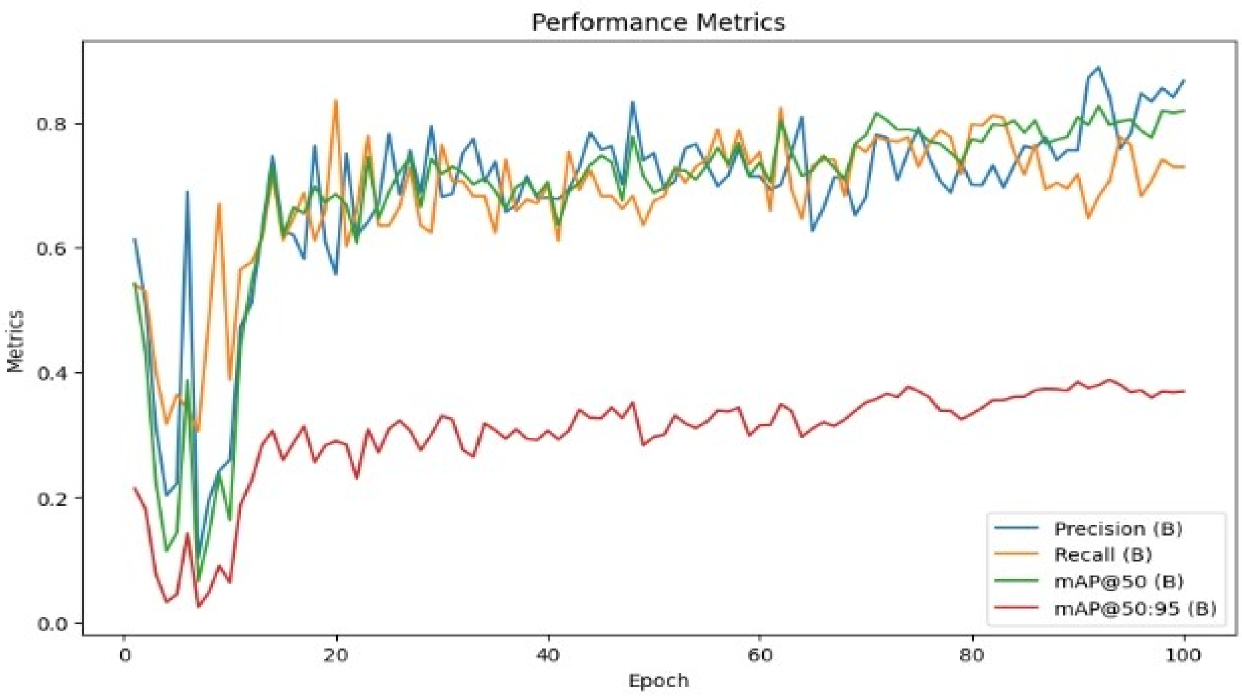

Figure 3: Performance Metrics Graph

The graph shows the bird detection model's performance over 100 epochs. **Precision** steadily improves, stabilizing at 0.85, indicating strong accuracy in positive detections. **Recall** increases initially but stabilizes at 0.7, suggesting the model misses some true positives. **mAP@50** stabilizes at 0.8, demonstrating good performance in simpler detection scenarios. However, **mAP@50:95** stabilizes at a lower value of 0.39, reflecting challenges in accurate localization for more complex cases. The model demonstrates strong precision and excels in simpler detection, but struggles with recall and localization in difficult scenarios, showing stable performance throughout training.

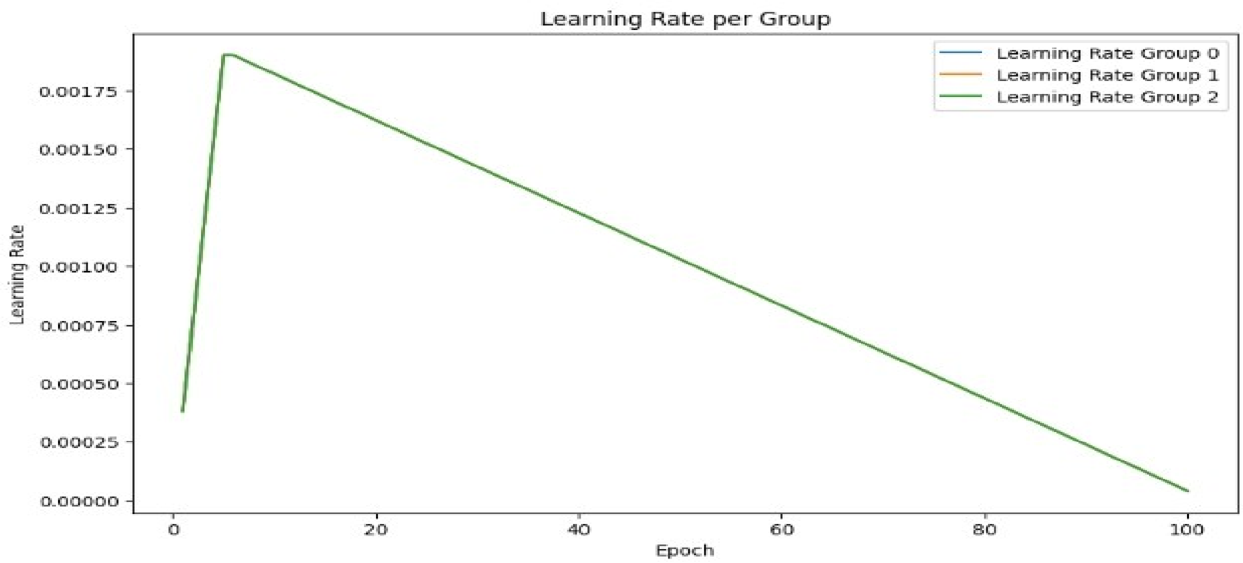

Figure 4: Learning Rate Per Group

The graph depicts the learning rate schedule during training, with a sharp initial rise that peaks around epoch 5. This initial **warm-up phase** allows the model to adapt efficiently to the data and prevents destabilizing large parameter updates. After reaching its peak, the learning rate gradually declines linearly over the remaining epochs, ensuring a smooth convergence towards optimal parameters and avoiding overshooting. The schedule is consistent across three parameter groups (**Learning Rate Group 0, Group 1, and Group 2**), indicating uniform treatment of different parameter sets. Overall, the learning rate schedule effectively balances **rapid initial learning** with **stable convergence**, facilitating efficient training.

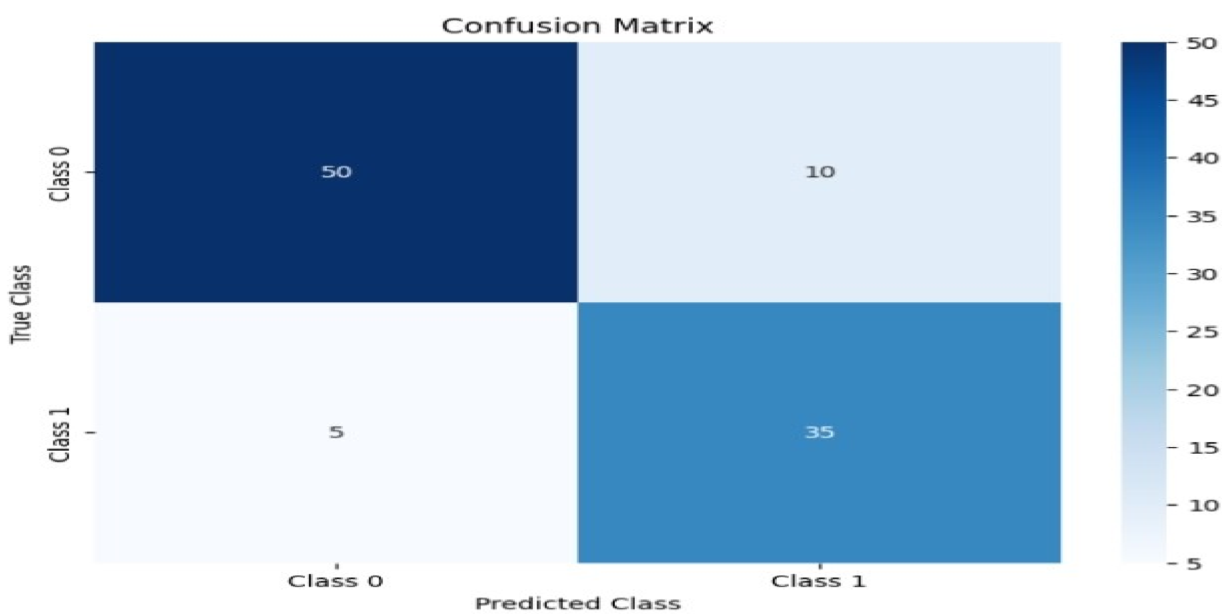

Figure 5: confusion Matrix

This confusion matrix provides an overview of the model's classification performance across two categories. The model accurately identified 35 instances of Class 1 and 50 instances of Class 0. However, 10 instances of Class 0 were misclassified as Class 1, while 5 instances of Class 1 were incorrectly labeled as Class 0. The model exhibits high recall for Class 1, indicating its effectiveness in identifying most instances of this category. However, the relatively lower precision suggests some misclassification of Class 0 as Class 1. Based on the confusion matrix, the model achieved an accuracy of 85%, a precision of 77%, a recall of 87%, and an F1-score of 82%.

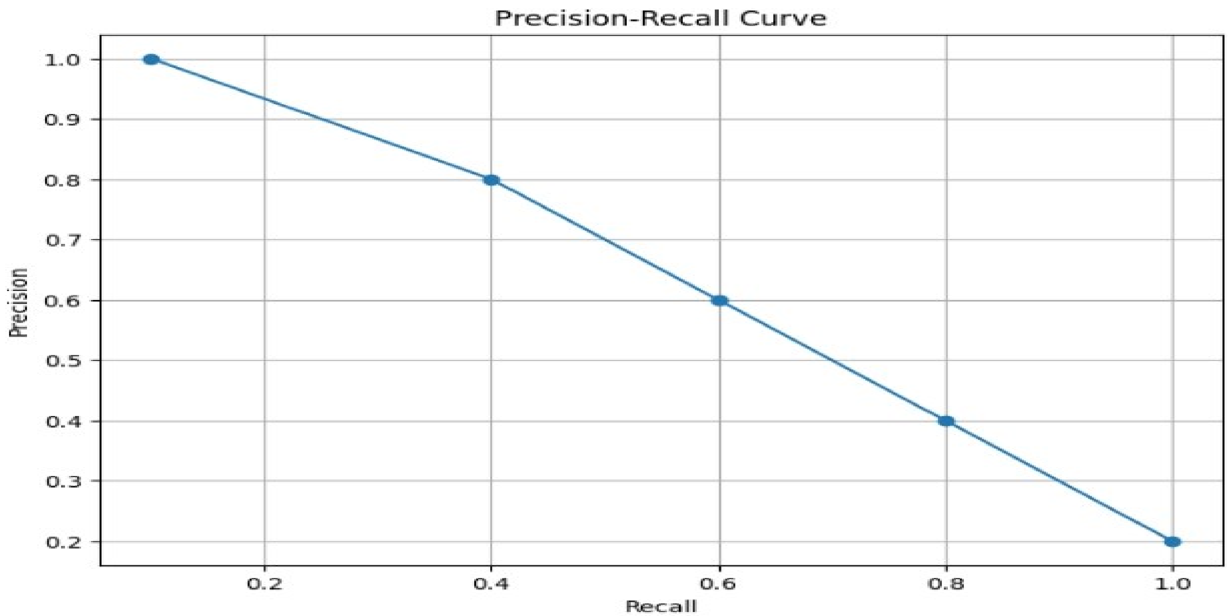

Figure 6: Precision-recall Curve

This precision-recall curve illustrates the balance between precision and recall across different classification thresholds. On the left side of the curve, where recall is low, precision remains high, indicating strong confidence in the model’s predictions. Conversely, on the right side, as recall increases, precision declines, suggesting a trade-off between capturing more positive instances and maintaining prediction accuracy. This indicates the model captures most of the positive cases. The curve's shape demonstrates how well the model maintains precision as recall increases. A more balanced precision-recall trade-off is typically preferred, depending on the use case.

Real-time inference

The model was exported to ONNX format and deployed on the microprocessor. We used a Raspberry Pi 4 Model B, selected for its balance of processing power and energy efficiency. The board was configured with a Raspberry Pi Camera Module v2 to capture video frames for inference. We tested the model on completely unseen images and real-time videos. We simulated different real-world scenarios, including variations in lighting and rotated bird views. Below is a description of some of the results we found:

Figure 7: Test Image1

The above image shows a bird correctly detected with a bounding box and a confidence score of **91%**. The detection is accurate, with the bounding box closely aligned to the bird's position. A confidence score of 91% signifies the model's high certainty in its prediction, demonstrating effective feature learning. This high-confidence detection suggests the robust performance of the model under normal lighting and simple backgrounds.

Figure 8: Test Image2

For the above image, the model successfully identifies both birds with high confidence (**0.90**), which indicates robust performance in this scenario. The background is cluttered with vegetation, making it harder to distinguish objects. Despite this, the model demonstrates effective generalization and avoids false positives. High confidence scores indicate the model's reliability in identifying small birds in challenging environments. The absence of overlapping or redundant boxes ensures precise detection.

Figure 9: Test Image3

The image depicts a dense flock of birds flying over a field, with bounding boxes and confidence scores showing detected birds. The model successfully identifies many birds, with confidence scores ranging from **0.53 to 0.86**. However, some birds are not detected due to **occlusion** or blending with the background. Lower confidence scores (e.g., **0.53** and **0.58**) are observed in areas with partial visibility or when birds are smaller and farther away. Despite the crowd density, the model performs well in detecting birds in moderately clear areas. Overall, the bird detection model demonstrates strong performance in identifying birds under optimal conditions but faces challenges in highly **crowded**, **occluded**, or **small-object-intensive** scenarios. The model's performance indicates robustness in clear conditions but highlights areas for improvement in handling complex scenes with overlapping birds.

CONCLUSION

This study presents a pioneering approach to mitigating bird-related crop losses in Nigerian rice farming by integrating UAVs with advanced computer vision models. The YOLOv8-based system achieved reliable detection and deterrence capabilities, addressing critical agricultural challenges. Although the model performs well under standard conditions, its effectiveness declines in densely populated and occluded environments, indicating areas for improvement. Future research should prioritize refining detection algorithms and integrating collaborative UAV systems to address complex scenarios. This study highlights the significant impact of AI-powered agricultural solutions in enhancing food security and fostering economic stability.

CONFLICT-OF-INTEREST DISCLOSURE

The authors confirm that there are no conflicts of interest associated with this research.

ACKNOWLEDGMENTS

We express our sincere gratitude to the Federal Ministry of Communication, Innovation, and Digital Economy, and the National Information Technology Development Agency (NITDA) for their generous support. This research was funded under the Nigerian Artificial Intelligence Research Scheme (NAIRS 2024), grant number: NITDA/HQ/RG/AI3361090986.

REFERENCES

Acharya, D., Saqib, M., Devine, C., Untiedt, C., Little, L. R., Wang, D., & Tuck, G. N. (2024). Using deep learning to automate the detection of bird-scaring lines on fishing vessels. *Biological Conservation*, 296(November 2023), 110713. https://doi.org/10.1016/j.biocon.2024.110713

Agha, S., Mohsan, H., Khan, M. A., Qasem, N., Othman, H., Li, Y., & Alsharif, M. H. (2023). Unmanned aerial vehicles (UAVs): practical aspects , applications , open challenges , security issues , and future trends. Intelligent Service Robotics, **16**(1): 109–137. https://doi.org/10.1007/s11370-022-00452-4

Agossou, H., Assede, E. P. S., Dossou, P. J., & Biaou, S. H. (2022). Effect of bird scaring methods on crop productivity and avian diversity conservation in agroecosystems of Benin. *International Journal of Biological and Chemical Sciences*, **16**(2): 527–542. https://doi.org/10.4314/ijbcs.v16i2.2

Bhusal, S., Karkee, M., Bhattarai, U., Majeed, Y., & Zhang, Q. (2022). Automated execution of a pest bird deterrence system using a programmable unmanned aerial vehicle (UAV).

Coluccia, A., Fascista, A., Schumann, A., Sommer, L., Dimou, A., Zarpalas, D., Méndez, M., de la Iglesia, D., González, I., Mercier, J. P., Gagné, G., Mitra, A., & Rajashekar, S. (2021). Drone vs. Bird detection: Deep learning algorithms and results from a grand challenge. *Sensors*, **21**(8): 1–27. https://doi.org/10.3390/s21082824

Kale, M., Balfors, B., Mörtberg, U., Bhattacharya, P., & Chakane, S. (2012). Damage to agricultural yield due to farmland birds, present repelling techniques and its impacts : an insight from the Indian perspective Sanctuaries in India protecting Bird Biodiversity which are Wild Ass. **8**(1): 49–62.

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 147, 70-90. https://doi.org/10.1016/j.compag.2018.02.016

Krizhevsky, B. A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. https://doi.org/10.1145/3065386

Liu, Z., & Li, J. (2023). Application of Unmanned Aerial Vehicles in Precision Agriculture. Marcoň, P., Janoušek, J., Pokorný, J., Novotný, J., Hutová, E. V., Širůčková, A., Čáp, M., Lázničková, J., Kadlec, R., Raichl, P., Dohnal, P., Steinbauer, M., & Gescheidtová, E. (2021). A system using artificial intelligence to detect and scare bird flocks in the protection of ripening fruit. Sensors, 21(12). https://doi.org/10.3390/s21124244

Mey, Y. De, & Demont, M. (2014). Bird Damage to Rice in Africa: October. https://doi.org/10.1079/9781845938123.0241 Mujahid, M., Rustam, F., Shafique, R., Montero, E. C., Alvarado, E. S., De, I., Diez, T., & Ashraf, I. (2024). Efficient deep learning - based approach for malaria detection using red blood cell smears. *Scientific Reports*, 1–16. https://doi.org/10.1038/s41598-024-63831-0

Okoli, V., & Ezui, G. (2012). Bird Scaring Technologies in Rice Production: The need for Policies Prohibiting Participation of Women and Chil ... Bird Scaring Technologies in Rice Production: The need for Policies Prohibiting Participation of Women and Children. *Accelerating the World's Research*, **1**(3): 031–038.

Rabiu, L., Ahmad, A., & Gohari, A. (2024). Advancements of Unmanned Aerial Vehicle Technology in the Realm of Applied Sciences and Engineering: A Review Advancements of Unmanned Aerial Vehicle Technology in the Realm of Applied Sciences and Engineering: A Review. March. https://doi.org/10.37934/araset.40.2.7495

Riya, R., KR, V., Sonamsi, S., & Jain, D. (2020). Automated Bird Detection and Repeller System Using IOT Devices: An Insight from Indian Agriculture Perspective. *SSRN Electronic Journal*, 3–5. https://doi.org/10.2139/ssrn.3563395

Roihan, A., Hasanudin, M., & Sunandar, E. (2020). Evaluation Methods of Bird Repellent Devices in Optimizing Crop Production in Agriculture. *Journal of Physics: Conference Series*, **1477**(3): 0–7. https://doi.org/10.1088/1742- 6596/1477/3/032012

Suleiman, I., Federal, T., Bida, P., Babawuya, A., & Salihu, B. A. (2021). A Review of Bird Pest Repellent Systems in Farms A Review of Bird Pest Repellent Systems in Farms. November.

Tetila, E. C., Machado, B. B., Astolfi, G., Belete, N. A. D. S., Amorim, W. P., Roel, A. R., & Pistori, H. (2020). Detection and classification of soybean pests using deep learning with UAV images. Computers and Electronics in Agriculture, 179, 105836. https://doi.org/10.1016/j.compag.2020.105836

Wang, H., Xu, Y., Yu, Y., & Lin, Y. (2022). An Efficient Model for a Vast Number of Bird Species Identification Based on Acoustic Features.

Wang, Z. (2021). Intelligent UAVs for Pest Bird Control in Vineyards. Wiafe, E. D. (2001). An inventory of birds and extent of damage to rice farms at the Kpong Irrigation Dam in the Lower Volta B asin, Ghana. **58**(1): 1–11.

Yao, G., Zhu, S., Zhang, L., & Qi, M. (2024). Hp-yolov8: high-precision small object detection algorithm for remote sensing images. *Sensors*, **24**(15): 4858.

Zhang, C. and Kovacs, J.M. (2012) The Applications of Small Unmanned Aerial Systems for Precision Agriculture: A Review. *Precision Agriculture*, **13**: 693-712. https://doi.org/10.1007/s11119-012-9274-5