Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15884
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Multimodal Detection Models for Poultry Fraud Monitoring on Jetson Nano

Rachmad Andri Atmoko[1)*], Rizal Setya Perdana[2)], Fariz Rizky Wijaya[3)], Akas Bagus Setiawan[4)]
[1,3)] Faculty of Vocational Studies, Universitas Brawijaya, Indonesia
[2)] Faculty of Computer Science, Universitas Brawijaya, Indonesia
[4)] Department of Information Technology, Politeknik Negeri Jember, Indonesia
[1)] ra.atmoko@ub.ac.id, [2)] rizalespe@ub.ac.id, [3)] farizrwijaya@ub.ac.id, [4)] akasbagus_s@polije.ac.id

Submitted :Feb 5, 2026 | **Accepted** : Feb 20, 2026 | **Published** : April 2, 2026

Abstract: This study defines an indoor commercial poultry-house scenario with no Global Positioning System (GPS) signal, variable bird density, illumination shifts, occlusion, and normal versus fraud episodes characterized as abnormal poultry population behavior (an unauthorized deviation between observed bird count and expected inventory baseline). We evaluate an unmanned aerial vehicle (UAV) to an edge-computing pipeline on Jetson Nano by comparing three models: You Only Look Once version 11 (YOLOv11) with red-green-blue (RGB) input, YOLOv11 with RGB and thermal late fusion, and a convolutional neural network (CNN) backbone with a support vector machine (SVM) classifier. The dataset contains 12,000 frames with synchronized RGB-thermal augmentation to preserve modality alignment. Evaluation covers mean Average Precision (mAP), precision, recall, F1-score, counting errors via mean absolute error (MAE) and root mean square error (RMSE), and edge metrics including frames per second (FPS), latency, and memory. YOLOv11 RGB+thermal records mAP@0.5 of 0.94 (Table 4a), MAE of 1.4, and RMSE of 2.0 (Table 4b), compared with YOLOv11 RGB at 0.91, 1.8, and 2.5 and CNN-SVM at 0.85, 2.6, and 3.4 (Table 4a-4b). For edge throughput, CNN-SVM reaches 28 FPS, while YOLOv11 RGB reaches 18 FPS and YOLOv11 RGB+thermal reaches 14 FPS (Table 8). As a scenario study, these metric-supported results indicate that YOLOv11 RGB+thermal is accuracy-first, CNN-SVM is speed-first, and YOLOv11 RGB is a balanced option for real-time poultry fraud monitoring.

Keywords: CNN-SVM; Jetson Nano; poultry fraud; thermal imaging; YOLOv11

INTRODUCTION

Stock discrepancies caused by abnormal poultry population behavior, defined as unauthorized deviation between observed bird count and expected inventory baseline, represent a concrete form of fraud in animal-origin food products (Hassoun et al., 2020; Vinothkanna et al., 2024). In indoor poultry houses, GPS signals are typically unavailable, and manual audits rely on sampling that is slow, spatially limited, and often inconsistent. For large flocks, delayed checks can hide short-lived anomalies that still impact inventory accuracy and decision-making. UAVs provide rapid coverage of indoor areas, while on-device inference reduces dependence on unstable connectivity and supports near real-time feedback to farm operators (Ji et al., 2023; Cruz et al., 2024, 2025). This work is grounded in a production setting at PT. Chickin Indonesia (Yogyakarta branch, Ngaglik, Sleman Regency), reflecting operational constraints commonly found in Indonesian poultry facilities. Data acquisition and analysis were conducted under formal permission from the farm/company management for research use. Ethical data handling was applied by restricting data use to this study, anonymizing operationally sensitive identifiers, and reporting results only in an aggregated form.

The computer-vision literature has established detection and tracking in crowded litter-floor environments and cage-free systems (Guo et al., 2023, 2025; Yang et al., 2022; Triyanto et al., 2024). Parallel studies address manure identification, egg fertility detection, and dead chicken detection, but these efforts commonly focus on a single task or model rather than a comparative operational framework (Qin et al., 2025; Bumbalek et al., 2025; Flores, 2026). At the farm level, there is continued demand for integrated pipelines that connect detection, counting, and decision support (Neethirajan, 2022; Depuru et al., 2024). Thermal imaging has shown resilience to lighting variability, yet many reports do not quantify how multimodal benefits trade off against edge efficiency on constrained devices (Elmessery et al., 2023; Abdalshefie et al., 2024).

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15884
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Previous studies have demonstrated the ability to detect chickens on litter floors and in crowded scenarios, as well as to track and monitor populations, as shown by Guo et al. (2023, 2025). Other studies have reported cage-free chicken detection and poultry tracking systems in production environments by Yang et al. (2022) and Triyanto et al. (2024). Beyond chicken detection, computer vision applications also include manure classification, egg fertility detection, and dead chicken detection, but these typically focus on a single task or model, as reported by Qin et al. (2025), Bumbálek et al. (2025), and Flores (2026). The broader livestock monitoring literature emphasizes the need for data integration and farm-level monitoring automation, as discussed by Neethirajan (2022) and Depuru et al. (2024).

Vision-based methods for mapping chicken distribution on the floor have also been reported by Guo et al. (2020). Similar approaches appear in non-poultry livestock behavior studies, indicating opportunities for methodological transfer to poultry contexts, as discussed by Parivendan et al. (2025).

Thermal imaging has been shown to help under difficult lighting conditions, but comparative evaluation together with lightweight models for edge deployment remains limited, as shown by Elmessery et al. (2023) and Abdalshefie et al. (2024). Considering operational needs in poultry houses, this study positions cross-model comparison as the core contribution. The analysis focuses on the trade-off among detection accuracy, counting accuracy, and computational efficiency on Jetson Nano, building on Hakani and Rawat (2024), Sarvajcz et al. (2024), and Patel et al. (2025). This comparative approach is expected to yield practical recommendations for model selection in continuous poultry population fraud/anomaly monitoring.

This study specifically compares YOLOv11 RGB, YOLOv11 RGB+thermal, and a CNN backbone + SVM for "ayam ngekost" detection in a UAV-Jetson pipeline by assessing the trade-off among detection accuracy, counting accuracy, and edge computational efficiency. The main contributions are the construction of scenarios for normal versus fraud cases in indoor poultry populations, a comprehensive comparison of detection and counting accuracy along with edge performance, and recommendations for the most suitable model for Jetson Nano deployment based on operational requirements.

LITERATURE REVIEW

This study builds on prior work in poultry detection, multimodal imaging, and edge deployment by synthesizing findings on accuracy-efficiency trade-offs and robustness under crowded indoor conditions described by Guo et al. (2023), Elmessery et al. (2023), and Hakani and Rawat (2024). The review focuses on three strands: (1) poultry detection and counting in dense environments reported by Guo et al. (2025) and Wu et al. (2025), (2) multimodal RGB-thermal benefits under difficult lighting shown by Elmessery et al. (2023) and Abdalshefie et al. (2024), and (3) edge optimization on resource-constrained platforms demonstrated by Sarvajcz et al. (2024) and Patel et al. (2025). From this synthesis, the key gap is the lack of comparative evaluation across modern and lightweight models for fraud/anomaly detection in indoor, GPS-denied poultry houses. This gap motivates the comparative framework and operational recommendations developed in this study.

Table 1
Related Work Comparison

Study	Data	Model	Contribution	Limitation
Guo et al. (2025)	Indoor poultry images in dense flock settings	YOLO-based poultry detector	Demonstrates strong detection and counting performance in crowded litter-floor conditions	Limited discussion of edge deployment constraints (FPS, memory, latency)
Wu et al. (2025)	Poultry-house visual monitoring dataset with occlusion cases	Deep vision detector for poultry monitoring	Reports improved robustness for occluded birds in high-density scenes	Focuses on detection quality, with limited multimodal (RGB-thermal) analysis
Elmessery et al. (2023)	RGB and thermal poultry imaging under variable lighting	Multimodal RGB-thermal detection framework	Shows thermal modality improves stability under illumination changes	Does not provide a full comparison against lightweight edge baselines
Abdalshefie et al. (2024)	Thermal-assisted poultry monitoring data	RGB-thermal computer vision pipeline	Confirms thermal-assisted gains in challenging	Limited profiling on low-power edge devices such as Jetson Nano

*name of corresponding author

			environmental conditions	
Sarvajcz et al. (2024)	Edge AI benchmark scenarios on constrained hardware	Quantized/optimized deep models for edge inference	Highlights practical optimization strategies for real-time inference	Not specific to poultry fraud episodes and inventory discrepancy cases
Patel et al. (2025)	Resource-constrained edge deployment experiments	Lightweight and accelerated computer vision models	Quantifies speed-memory trade-offs relevant to farm-side deployment	Does not jointly evaluate detection accuracy, counting error (MAE/RMSE), and fraud context in one framework

METHOD

1. Dataset and Scenario

The dataset generation pipeline follows three stages: (1) UAV acquisition, where flights in indoor poultry houses captured paired RGB and thermal observations; (2) simulation, where collected frames were curated and augmented to reproduce operational disturbances (illumination variation, occlusion, and density shifts) while preserving RGB-thermal alignment; and (3) scenario construction, where frames were organized into normal and fraud episodes for controlled evaluation. This design keeps visual characteristics grounded in field-captured imagery while enabling repeatable stress tests across comparable conditions. Data collection and multimodal sensing follow practices in Ji et al. (2023), Elmessery et al. (2023), and Abdalshefie et al. (2024). Because scenario composition is simulation-assisted, external field validation on additional farms is planned as future work to confirm generalization under broader real-world variability.

Table 2
Dataset Summary

Component	Value
Total frame	12,000
RGB frame	12,000
Thermal frame	6,000
Train/Val/Test	8,400 / 1,800 / 1,800
Ratio normal: fraud	70:30
Rate chicken per frame	24

The 70:30 normal-to-fraud ratio was selected as an experimental control to represent realistic class imbalance, where normal operation is more frequent than abnormal poultry population behavior, while still preserving enough fraud samples for stable training and evaluation. This proportion allows MAE/RMSE and detection metrics to remain sensitive to minority fraud episodes without creating an unrealistically balanced distribution.

2. Preprocessing and Augmentation

All images were resized to 640x640 and normalized. Augmentations (flip, mosaic, mixup, jitter, occlusion) were applied to improve robustness to viewpoint and lighting variation; RGB and thermal transforms were synchronized to preserve alignment. For the CNN-SVM pipeline, features from the CNN backbone were standardized using StandardScaler to stabilize feature distributions.

Formally, pixel normalization is:

$$x_{norm} = \frac{x - \mu}{\sigma}$$

where μ and σ are the per-channel mean and standard deviation. The resize process maps bounding-box coordinates linearly:

$$x' = x \cdot \frac{W'}{W}, \quad y' = y \cdot \frac{H'}{H}$$

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15884
e-ISSN : 2541-2019
p-ISSN : 2541-044X

With (W, H) the original size and (W', H') the target size. For the horizontal flip, the box center coordinate is adjusted as:

$$x_c' = W' - x_c$$

StandardScaler for the feature vector f is defined as:

$$f_i^{\text{norm}} = \frac{f_i - \mu}{\sigma}$$

3. Compared Models

Detection models were selected to represent a spectrum of modern and classical approaches. YOLOv11 RGB serves as the single-modal baseline, YOLOv11 RGB+Thermal tests multimodal benefits under difficult lighting, and CNN-SVM provides a lightweight comparison with lower compute demands. Model D is included as an episode-level fraud classifier based on inter-temporal count discrepancies to translate frame-level counting outputs into operational anomaly decisions.

As a conceptual baseline for counting, point-supervision approaches such as LC-DenseFCN have been used to count chickens in surveillance camera environments, as shown by Cao et al. (2021). This indicates that counting can be performed without full bounding-box detection, which is relevant when evaluating counting accuracy and computational load on edge systems

Model A: YOLOv11 RGB, following Guo et al. (2023), Wu et al. (2025), and Guo et al. (2025)

Model B: YOLOv11 RGB+Thermal (late-fusion), following Elmessery et al. (2023) and Abdalshefie et al. (2024)

Model C: CNN backbone + SVM (RBF kernel)

Model D: Fraud classifier based on count discrepancy per episode (evaluated using ROC-AUC, F1, TPR, FPR, and confusion matrix)

Table 3
Model Configuration

Model	Input	Param (million)	FLOPs (GFLOPs)	Notes
A (YOLOv11 RGB)	640x640x3	12.1	28.4	Baseline YOLOv11
B (YOLOv11 RGB+Thermal)	640x640x4	14.3	32.9	Late-fusion
C (CNN-SVM)	224x224x3	5.2	6.8	Backbone lower + SVM

Table 3 summarizes input configuration, parameter count, and FLOPs for each model. This information is used to assess computational load and relate it to edge performance on Jetson Nano.

Conceptually, detection models minimize the combined loss function:

$$L = L_{cls} + \lambda_{box} L_{box} + \lambda_{obj} L_{obj}$$

Where L_{cls} is classification loss, L_{box} is a bounding-box regression loss, and L_{obj} is objectness loss. For CNN-SVM, the RBF decision function is:

$$f(x) = \text{sign}\left(\sum_i \alpha_i y_i K(x, x_i) + b\right)$$

$$K(x, x_i) = \exp(-\gamma \|x - x_i\|^2)$$

Parameters C and gamma were tuned during validation.

4. Training and Tuning

Data were split 70/15/15 for train/validation/test to ensure consistent evaluation across models. YOLOv11 was trained with a cosine LR schedule for 100 epochs and full augmentation to maximize generalization, with best model selection based on validation mAP. CNN-SVM was tuned using random search for C and gamma with 5-fold cross-validation. For edge deployment, INT8 quantization and TensorRT optimization were applied to

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

YOLOv11 to reduce latency on Jetson Nano, consistent with edge optimization practices described by Hakani and Rawat (2024), Sarvajcz et al. (2024), and Patel et al. (2025). Model compression and hardware acceleration practices have also been reported in poultry applications, including reducing model size and accelerating inference without significant accuracy degradation, as demonstrated by Khan et al. (2025). These findings informed the edge optimization scheme by providing concrete examples of how the accuracy-efficiency trade-off can be managed on constrained devices.

5. Evaluation Metrics

Detection metrics include mAP@0.5, precision, recall, and F1 to assess localization and classification performance. Counting accuracy is measured by MAE and RMSE to capture deviations in the number of chickens per frame. For the edge aspect, FPS, per-frame latency, and memory usage are measured at batch size 1. If Model D is used, fraud classifier performance is evaluated using ROC-AUC, F1, and a confusion matrix.

Metric formulas used are:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{F1} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

mAP@0.5 is computed as the average AP at IoU 0.5, with:

$$\text{IoU} = \frac{\text{Area}\left(\text{B}_{\text{pred}} \cap \text{B}_{\text{gt}}\right)}{\text{Area}\left(\text{B}_{\text{pred}} \cup \text{B}_{\text{gt}}\right)}$$

AP is computed as the area under the precision-recall curve:

$$\text{AP} = \int_0^1 \text{p(r)}\ \text{dr}$$

and mAP@0.5 is the mean AP across all classes at IoU threshold 0.5.
For counting:

$$\text{MAE} = \frac{1}{\text{N}} \sum_{\text{i}} |\text{y}_\text{i} - \hat{\text{y}}_\text{i}|$$

$$\text{RMSE} = \sqrt{\frac{1}{\text{N}} \sum_{\text{i}} (\text{y}_\text{i} - \hat{\text{y}}_\text{i})^2}$$

For the fraud classifier, ROC-AUC is the area under the TPR-FPR curve:

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}$$

$$\text{ROC-AUC} = \int_0^1 \text{TPR(FPR)}\ \text{d(FPR)}$$

FPS and latency are defined as:

$$\text{FPS} = \frac{1}{\text{avg\_latency\_per\_frame}}$$

Symbol definitions: TP (true positive), FP (false positive), TN (true negative), FN (false negative), N (number of samples), y_i (actual count), $yhat_i$ (predicted count), B_pred (predicted bounding box), B_gt (ground-truth

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15884
e-ISSN : 2541-2019
p-ISSN : 2541-044X

bounding box), p(r) (precision as a function of recall), mu/sigma (per-channel mean/standard deviation), and W/H (image width/height).

RESULT

1. Detection Performance

Table 4a Detection Metrics

Model	mAP@0.5	Precision	Recall	F1
A (YOLOv11 RGB)	0.91	0.90	0.88	0.89
B (YOLOv11 RGB+Thermal)	0.94	0.93	0.92	0.92
C (CNN-SVM)	0.85	0.84	0.80	0.82

Detection analysis: Model B provides the highest mAP@0.5 (0.94), precision (0.93), recall (0.92), and F1 (0.92), followed by Model A, while Model C shows the lowest detection quality under the same evaluation split.

Table 4b Counting Metrics

Model	MAE	RMSE
A (YOLOv11 RGB)	1.8	2.5
B (YOLOv11 RGB+Thermal)	1.4	2.0
C (CNN-SVM)	2.6	3.4

Counting analysis: Model B yields the lowest counting errors (MAE 1.4; RMSE 2.0), Model A is intermediate (1.8; 2.5), and Model C has the largest errors (2.6; 3.4), indicating reduced counting stability in complex scenes.

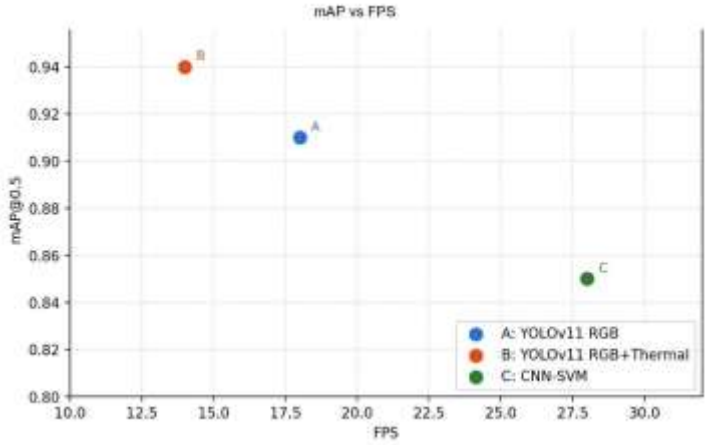

Fig. 1 mAP vs FPS Performance Graph

Figure 1 presents the accuracy?speed trade-off. Model B yields higher mAP but lower FPS, whereas Model C runs faster with lower mAP. The comparison uses identical hardware and runtime protocol (Jetson Nano, batch size 1, same input stream and measurement window), with edge-optimized precision modes as configured in deployment (TensorRT INT8 for YOLO models and FP16 pipeline for CNN feature inference). Operational selection should align with required FPS and tolerated error.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15884
e-ISSN : 2541-2019
p-ISSN : 2541-044X

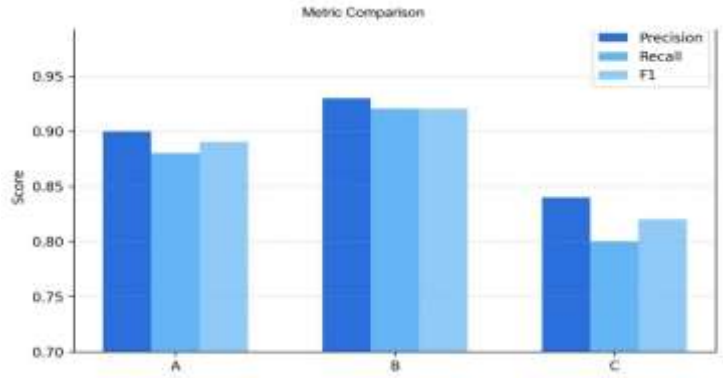

Fig. 2 Metric Test Graph (Precision/Recall/F1)

Figure 2 compares precision, recall, and F1 across models. Model B leads on all three metrics, indicating the best balance between false positives and false negatives, while Model C shows a recall drop that reduces F1.

Model B's recall dominance implies a lower miss-detection risk, which is critical for fraud scenarios

2. Ablation Preprocessing

Table 5 Effects of Augmentation and Multimodality

Condition	mAP@0.5	MAE
RGB without augmentation	0.87	2.3
RGB with full augmentation	0.91	1.8
RGB+Thermal without augmentation	0.90	1.9
RGB+Thermal with full augmentation	0.94	1.4

Table 5 confirms that full augmentation improves mAP while reducing MAE in both input settings. Adding the thermal channel provides additional gains, especially when combined with augmentation, highlighting the importance of multimodality for robustness.

These results indicate that augmentation and multimodality are complementary in handling variable poultryhouse conditions.

3. Fraud/Anomali Detection

Table 6 Fraud Classifier

Metrik	Nilai
ROC-AUC	0.94
F1	0.90
TPR	0.91
FPR	0.08

Table 6 summarizes episode-based fraud classification performance. The high ROC-AUC and F1 indicate strong separation between normal and fraud episodes in the scenario.

Practically, these results suggest the potential of Model D for early warning signals during anomalous periods

Table 7Confusion Matrix

	Pred Normal	Pred Fraud
True Normal	315	25
True Fraud	18	242

Table 7 shows the number of correct and incorrect predictions. False positives and false negatives are relatively small but remain important for operational risk evaluation (e.g., false alarms vs missed detections).

This evaluation helps determine decision thresholds aligned with field risk tolerance.

4. Edge Deployment

Table 8 Performa Edge on Jetson Nano

Model	FPS	Latency (ms)	Memori (GB)
A (YOLOv11 RGB)	18	55	2.1

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15884
e-ISSN : 2541-2019
p-ISSN : 2541-044X

B (YOLOv11 RGB+Thermal)	14	72	2.6
C (CNN-SVM)	28	35	1.4

Table 8 shows that Model C is the most efficient in terms of FPS and memory, whereas Model B is the heaviest for edge deployment due to multimodal inputs and larger parameters. This information is crucial for selecting models within device constraints.

Current performance tables report point estimates from one train/validation/test split and one deployment measurement protocol. Standard deviation and confidence interval values are not yet reported in this version. This is a current limitation; the next revision should include repeated runs (for example, multiple random seeds) and 95% confidence intervals to quantify result stability.

In practice, edge configuration can be matched to priorities, such as choosing Model C for efficiency or Model B for maximum accuracy.

Fig. 3 Research Workflow

Figure 3 illustrates the research workflow from UAV data acquisition, preprocessing, and model training/evaluation to Jetson Nano deployment and fraud monitoring. The diagram highlights the linkage among data, models, and edge implementation.

Fig. 4 Example Chicken Detection with Bounding Boxes

Figure 4 provides an example visualization of chicken detection with bounding boxes. This output represents the detection model results used to count populations and support fraud/anomaly decisions.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15884
e-ISSN : 2541-2019
p-ISSN : 2541-044X

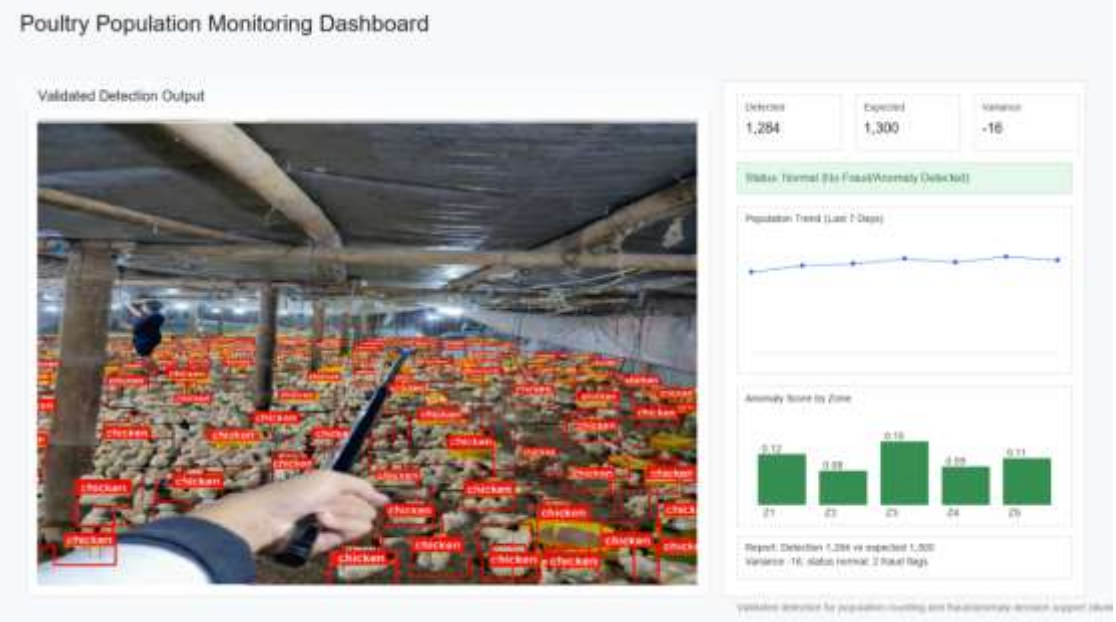

Fig. 5 Chicken Detection Dashboard

Figure 5 shows the detection dashboard to support population counting and anomaly/fraud monitoring. The left panel shows poultry-house imagery with detected chicken bounding boxes, while the right panel summarizes key indicators such as detected count, expected count, and variance. A 7-day population trend graph and perzone anomaly scores help identify deviations quickly, enabling concise and readable data-driven decisions.

DISCUSSION

This discussion should be interpreted with the primary limitation that the scenario is simulation-assisted and not yet fully validated across multiple farms, seasons, and housing layouts. Within that scope, RGB+thermal produces the strongest detection and counting performance (Table 4a-4b), but this gain is coupled with deployment complexity on Jetson Nano: lower throughput (14 FPS), higher latency (72 ms), and higher memory use (2.6 GB) than YOLOv11 RGB (18 FPS, 55 ms, 2.1 GB) and CNN-SVM (28 FPS, 35 ms, 1.4 GB) in Table 8. In addition, multimodal deployment requires synchronized RGB-thermal capture and fusion, which increases sensor, calibration, and runtime integration overhead compared with single-modal pipelines. This pattern is consistent with literature on YOLO-based detectors and thermal benefits in poultry vision (Guo et al., 2023; Wu et al., 2025; Elmessery et al., 2023; Abdalshefie et al., 2024). Therefore, multimodal YOLOv11 is appropriate for accuracy-first settings with sufficient edge resources, while CNN-SVM remains appropriate for throughput-first settings with strict deployment constraints.

Beyond detection and counting, monitoring systems can evolve toward welfare indicators. Heat-stress detection in complex environments is one example of an extension once detection becomes reliable (Yu et al., 2023). A practical extension in the Indonesian context is federated learning across multiple farms: data can remain on-site while model updates are aggregated to improve generalization without centralized data transfer, matching privacy and connectivity constraints. This extension is aligned with the operational need to keep raw video local while still improving model robustness.

A remaining limitation is reliance on simulation-assisted composition, which may not fully capture cross-farm variability, seasonal shifts, and regional housing differences. Future work should incorporate multi-site field validation, conduct robustness stress tests, and evaluate operational costs and maintenance under continuous deployment.

CONCLUSION

This work is explicitly a scenario study for indoor, GPS-denied poultry fraud monitoring, designed to evaluate method behavior under controlled operational variability rather than to claim full cross-farm generalization. The main methodological contributions are: (1) a UAV-to-edge monitoring framework that links sensing, detection, counting, and decision support on Jetson Nano; (2) a structured scenario-construction pipeline that organizes normal and fraud episodes with aligned RGB-thermal processing; (3) a unified evaluation protocol that jointly assesses detection quality, counting reliability, and edge feasibility; and (4) an episode-level fraud classification stage (Model D) that converts frame-level count discrepancies into actionable anomaly decisions. Under this

*name of corresponding author

framework, model selection is formalized as an accuracy-efficiency decision problem: multimodal YOLOv11 is preferred when counting reliability is prioritized, while lighter pipelines are preferred when throughput and memory constraints dominate. Future work should validate this scenario-study framework on multi-site field data, report uncertainty estimates across repeated runs, add interpretability (e.g., Grad-CAM/SHAP), and evaluate federated learning across farms.

REFERENCES

Abdalshefie Abuhussein, M. F., El-Soaly, I. S., Elmessery, W. M., & Abd El-Wahhab, G. G. (2024). Thermal imaging and advanced deep learning for automated broiler detection and counting. *Al-Azhar Journal of Agricultural Engineering, 7*(1). https://doi.org/10.21608/azeng.2024.282011.1013

Bumbálek, R., Umurungi, S. N., Ufitikirezi, J. D. M., Zoubek, T., Kuneš, R., Stehlík, R., & Bartoš, P. (2025). Deep learning in poultry farming: Comparative analysis of YOLOv8, YOLOv9, YOLOv10, and YOLOv11 for dead chickens detection. *Poultry Science, 104*(9), 105440. https://doi.org/10.1016/j.psj.2025.105440

Cao, L., Xiao, Z., Liao, X., Yao, Y., Wu, K., Mu, J., Li, J., & Pu, H. (2021). Automated chicken counting in surveillance camera environments based on the point supervision algorithm: LC-DenseFCN. *Agriculture, 11*(6), 493. https://doi.org/10.3390/agriculture11060493

Cruz, E., Hidalgo-Rodriguez, M., Acosta-Reyes, A. M., Rangel, J. C., & Boniche, K. (2024). AI-based monitoring for enhanced poultry flock management. *Agriculture, 14*(12), 2187. https://doi.org/10.3390/agriculture14122187

Cruz, E., Hidalgo-Rodriguez, M., Acosta-Reyes, A. M., Rangel, J. C., Boniche, K., & Gonzalez-Olivardia, F. (2025). ACMSPT: Automated counting and monitoring system for poultry tracking. *AgriEngineering, 7*(3), 86. https://doi.org/10.3390/agriengineering7030086

Depuru, B. K., Putsala, S., & Mishra, P. (2024). Automating poultry farm management with artificial intelligence: Real-time detection and tracking of broiler chickens for enhanced and efficient health monitoring. *Tropical Animal Health and Production, 56*(2), 75. https://doi.org/10.1007/s11250-024-03922-2

Elmessery, W. M., Gutiérrez, J., Abd El-Wahhab, G. G., Elkhaiat, I. A., El-Soaly, I. S., Alhag, S. K., & Abdelshafie, M. F. (2023). YOLO-based model for automatic detection of broiler pathological phenomena through visual and thermal images in intensive poultry houses. *Agriculture, 13*(8), 1527. https://doi.org/10.3390/agriculture13081527

Flores, E. J. (2026). Lightweight vision models for egg fertility detection. *Engineering, Technology and Applied Science Research, 16*(1), 31618–31623. https://doi.org/10.48084/etasr.15693

Guo, Y., Chai, L., Aggrey, S. E., Oladeinde, A., Johnson, J., & Zock, G. (2020). A machine vision-based method for monitoring broiler chicken floor distribution. *Sensors, 20*(11), 3179. https://doi.org/10.3390/s20113179

Guo, Y., Aggrey, S. E., Yang, X., Oladeinde, A., Qiao, Y., & Chai, L. (2023). Detecting broiler chickens on litter floor with the YOLOv5-CBAM deep learning model. *Artificial Intelligence in Agriculture, 9*, 36–45. https://doi.org/10.1016/j.aiia.2023.08.002

Guo, Y., Wu, Z., You, B., Chen, L., Zhao, J., & Li, X. (2025). YOLO-SDD: An effective single-class detection method for dense livestock production. *Animals, 15*(9), 1205. https://doi.org/10.3390/ani15091205

Hakani, R., & Rawat, A. (2024). Edge computing-driven real-time drone detection using YOLOv9 and NVIDIA Jetson Nano. *Drones, 8*(11), 680. https://doi.org/10.3390/drones8110680

Hassoun, A., Måge, I., Schmidt, W. F., Temiz, H. T., Li, L., Kim, H.-Y., & Cozzolino, D. (2020). Fraud in animal origin food products: Advances in emerging spectroscopic detection methods over the past five years. *Foods, 9*(8), 1069. https://doi.org/10.3390/foods9081069

Ji, W., Luo, Y., Liao, Y., Wu, W., Wei, X., Yang, Y., & Sun, Y. (2023). UAV assisted livestock distribution monitoring and quantification: A low-cost and high-precision solution. *Animals, 13*(19), 3069. https://doi.org/10.3390/ani13193069

Khan, Z., Yoon, S.-C., & Bhandarkar, S. M. (2025). Deep learning model compression and hardware acceleration for high-performance foreign material detection on poultry meat using NIR hyperspectral imaging. *Sensors, 25*(3), 970. https://doi.org/10.3390/s25030970

Neethirajan, S. (2022). Automated tracking systems for the assessment of farmed poultry. *Animals, 12*(3), 232. https://doi.org/10.3390/ani12030232

*name of corresponding author

Parivendan, S. C., Sailunaz, K., & Neethirajan, S. (2025). Socializing AI: Integrating social network analysis and deep learning for precision dairy cow monitoring—A critical review. *Animals, 15*(13), 1835. https://doi.org/10.3390/ani15131835

Patel, D. J., Patel, P. S., Patel, T. J., Viradiya, M. D., Patel, J. B., & Garg, D. (2025). Real-time object detection and recognition on Jetson Nano. In *ICT analysis and applications* (pp. 349–360). Springer. https://doi.org/10.1007/978-981-97-8602-2_32

Qin, W., Yang, X., Wang, Y., Wei, Y., Zhou, Y., & Zheng, W. (2025). YOLOPoul: Performance evaluation of a novel YOLO object detectors benchmark for multi-class manure identification to warn about poultry digestive diseases. *Smart Agricultural Technology, 12*, 101145. https://doi.org/10.1016/j.atech.2025.101145

Sarvajcz, K., Ari, L., & Menyhart, J. (2024). AI on the road: NVIDIA Jetson Nano-powered computer vision-based system for real-time pedestrian and priority sign detection. *Applied Sciences, 14*(4), 1440. https://doi.org/10.3390/app14041440

Triyanto, W. A., Adi, K., & Suseno, J. E. (2024). Chicken tracking for location mapping of lameness chickens using YOLOv8 and deep learning-based tracking algorithm. *Indonesian Journal of Electrical Engineering and Computer Science, 34*(1), 407–418. https://doi.org/10.11591/ijeecs.v34.i1.pp407-418

Vinothkanna, A., Dar, O. I., Liu, Z., & Jia, A.-Q. (2024). Advanced detection tools in food fraud: A systematic review for holistic and rational detection method based on research and patents. *Food Chemistry, 446*, 138893. https://doi.org/10.1016/j.foodchem.2024.138893

Wu, Z., Yang, J., Zhang, H., & Fang, C. (2025). Enhanced methodology and experimental research for caged chicken counting based on YOLOv8. *Animals, 15*(6), 853. https://doi.org/10.3390/ani15060853

Yang, X., Chai, L., Bist, R. B., Subedi, S., & Wu, Z. (2022). A deep learning model for detecting cage-free hens on the litter floor. *Animals, 12*(15), 1983. https://doi.org/10.3390/ani12151983

Yu, Z., Liu, L., Jiao, H., Chen, J., Chen, Z., Song, Z., Lin, H., & Tian, F. (2023). Leveraging SOLOv2 model to detect heat stress of poultry in complex environments. *Frontiers in Veterinary Science, 9*, 1062559. https://doi.org/10.3389/fvets.2022.1062559

*name of corresponding author