

Machine Learning in Commercial Poultry Farming: Computer Vision-Assisted Broiler Count Verification During Manual Poultry Harvest Weighing

Engr. Aris R. Salomon, MBA, MSCS
img2data.tech
Las Piñas, Philippines
Email: ars@img2data.tech

Abstract—Commercial poultry harvesting in many Philippine farm settings still depends on manual bird handling, weighing, and record verification procedures, particularly where conveyor-based systems remain too costly for routine use. These methods place substantial demands on workers and may prolong the time broilers spend in a stressful handling environment. This study evaluates oriented bounding box (OBB) object detection models for harvest-stage monitoring in a practical weighing-in and count verification scenario. A camera positioned near the weighing station captures broiler images while a computer vision pipeline performs automated detection of broiler heads. OBB detection is used to better represent the irregular orientations and overlapping positions of birds during manual handling. The study compares YOLO-based OBB detectors in terms of detection accuracy, localization quality, and count-verification relevance under operational farm conditions. Held-out test results showed that the medium model achieved the strongest pure localization performance, while the small and nano models provided stronger count-verification behavior at the frame and event levels, respectively. Unlike prior work focused on feeding, drinking, camera placement, and other poultry monitoring scenarios, this study examines computer vision performance during harvest operations, where count verification, workflow visibility, and documentation are directly relevant to farm management. The proposed approach is intended to support workers during verification while giving farm owners and veterinarians better visibility into harvest activity. By improving operational efficiency, the system may also help shorten manual weighing and count verification workflows, with potential welfare benefits for broilers during harvest.

Index Terms—Computer vision, machine learning, broiler harvesting, count verification, oriented bounding boxes.

I. INTRODUCTION

Commercial poultry production increasingly depends on timely and reliable monitoring to support bird welfare, farm operations, and more structured counting and record-keeping. Recent work in precision livestock farming has shown that digital sensing, machine learning, and automated monitoring can strengthen welfare assessment and management decisions when deployed in ways that reflect actual farm conditions [1], [2]. In poultry systems, these tools are particularly valuable because routine observation remains labor-intensive, difficult

to standardize, and highly dependent on flock density, environmental conditions, and worker workload.

Computer vision has emerged as one of the more practical paths for poultry monitoring because it can operate continuously using relatively accessible camera infrastructure. From a computer science perspective, this problem belongs to automated object detection in cluttered scenes, where accuracy depends on how well the system handles occlusion, scale variation, irregular orientation, and motion while still maintaining inference efficiency suitable for field use. Recent studies have applied vision-based methods to welfare-related poultry behaviors such as foraging and piling, while other work has extended automated detection to lesion identification and broader edge artificial intelligence deployment in poultry settings [3]–[6]. These studies collectively show that automated visual monitoring is promising, but they also make clear that performance depends heavily on the scene being monitored, the operational objective, and the degree of crowding or overlap among birds.

Dense-scene detection remains one of the central challenges in poultry computer vision. Commercial flocks frequently contain overlapping birds, irregular postures, partial occlusion, and rapid motion, all of which complicate reliable counting and localization. This difficulty has motivated work on dense chicken counting networks, density-map methods, and other detection strategies designed for visually crowded poultry environments [7], [8]. The recent availability of real-farm detection datasets further reinforces the importance of evaluating models under authentic poultry-house conditions rather than relying only on simplified imagery [9]. Related studies in automated video analysis and broiler weighing also show that useful farm decisions depend not only on detection accuracy, but also on whether model outputs remain reliable enough to support operational tasks such as verification, handling, and documentation [4], [10].

For harvest operations, these issues become even more consequential. In many Philippine commercial farm settings, broiler harvesting still relies on manual bird handling, weighing, and record verification procedures because conveyor-based systems are not yet economically feasible for a large

share of farms. Manual catching and transport are recognized as welfare-critical stages in poultry production, since handling methods can affect stress, injury risk, and worker burden [11]–[13]. When weighing and count verification activities are also performed manually, the process may become slower, harder to verify consistently, and more demanding for both workers and birds. This creates a practical opportunity for computer vision systems that improve verification without requiring full automation.

The present study builds directly on two earlier scenario-specific studies by the author. The first examined broiler detection in feeding and drinking contexts, comparing standard axis-aligned bounding boxes and oriented bounding boxes in dense poultry scenes [14]. The second investigated how camera placement and distance affect Artificial Intelligence-based broiler monitoring performance in high-density houses [15]. Those studies established two important points: first, that poultry computer vision performance is strongly shaped by the specific farm scenario being observed; and second, that deployment choices such as annotation geometry and camera distance materially affect real-world usability. However, neither study addressed the harvest stage, where the monitoring objective shifts from general behavior observation toward counting support, weighing workflow assistance, and record verification.

This paper therefore focuses on a different and operationally critical scenario: broiler harvesting, specifically the weighing-in and count verification stage. In this setting, a camera positioned near the weighing station observes broilers during manual harvest handling, while a computer vision system performs broiler-head detection near the point of operation. Oriented bounding boxes are adopted because the birds appear in irregular orientations and frequently overlap while being handled near the weighing point. The study evaluates YOLO-based OBB detectors in terms of detection accuracy, localization quality, and count-verification reliability under conditions intended to reflect practical farm deployment.

The intended role of the proposed system is operational support. Faster count verification, more consistent documentation, and improved workflow visibility can strengthen harvest control for farm owners, provide veterinarians with better visibility into a stressful production stage, and give workers a practical verification tool during a demanding manual process. If such support shortens the time birds spend in manual weighing and record verification procedures, it may also contribute to better welfare conditions during one of the more stressful stages of broiler production. In that sense, the contribution of this work lies not only in model evaluation, but also in framing harvest-stage computer vision as a welfare-aware and operations-aware decision-support tool for commercial poultry farms.

The main contributions of this paper are as follows:

- We examine a harvest-specific poultry computer vision scenario centered on manual weighing-in and count verification, which differs from prior feeding, drinking, and camera-placement studies.
- We evaluate YOLO-based oriented bounding box detectors for broiler head detection in dense and irregular harvest-stage visual conditions.

- We assess the system using not only detection performance but also operationally relevant measures such as frame-level and event-level count-verification reliability.
- We position the proposed approach as a welfare-aware operational support tool intended to enhance harvest verification, documentation, and efficiency in commercial poultry farms.

II. RELATED WORK

The literature relevant to this study sits at the intersection of poultry welfare, farm operations, computer vision, and precision livestock farming. However, much of the available work emphasizes either poultry biology and production performance on one side, or technical detection performance on the other. For a harvest-stage study such as the present one, the more useful question is not simply whether Artificial Intelligence can detect birds, but whether visual monitoring can support real work in a stressful and time-sensitive farm activity. This section therefore reviews prior work through six themes: welfare and operational context, vision-based poultry monitoring, dense-scene counting and verification, oriented object detection and practical deployment considerations, scenario-specific prior studies, and the research gap around harvest-stage weighing and count verification.

A. Poultry Welfare, Observation, and Operational Context

Welfare-oriented poultry monitoring has increasingly moved beyond simple production output and toward a broader view of how birds experience commercial housing, handling, and management conditions. Recent precision livestock farming studies have emphasized that digital tools should be evaluated in relation to actual animal conditions rather than as isolated engineering systems [1], [2]. This framing is important for commercial operations because the people making decisions on farms are not only concerned with detection performance, but also with whether a tool helps them observe flock condition, manage labor, and reduce avoidable stress during critical production stages.

In broiler production, manual handling remains one of the most sensitive points from both a welfare and operations perspective. Manual catching and transportation have been identified as stages where bird stress, injury risk, worker fatigue, and workflow inconsistency can intersect [11], [12]. Field comparisons between manual and mechanical harvesting have likewise shown that harvesting method influences stress-related outcomes, reinforcing the importance of treating harvest as a welfare-relevant operational event rather than only a logistics step [13]. For farm owners, this stage affects timing, labor coordination, and record accuracy. For veterinarians, it affects welfare oversight and practical handling conditions. For workers, it represents a physically demanding process where speed and accuracy are both expected. These concerns make harvest operations a legitimate monitoring target, even when the system being proposed is not a full automation platform.

This matters in the Philippine setting, where many farms continue to rely on manual harvesting rather than conveyor-assisted systems. In such environments, a useful Artificial

Intelligence system is one that helps workers verify counts, helps managers maintain cleaner records, and helps reduce unnecessary delays in bird handling. That practical orientation is consistent with a welfare-aware view of precision livestock farming, where technology is introduced to support judgment and workflow [1].

B. Computer Vision in Poultry Monitoring

Computer vision has already shown value in poultry monitoring, especially for continuous observation tasks that are difficult to sustain manually. Prior work has used machine vision to detect foraging behavior, piling behavior, and other indicators related to activity, space use, or welfare in commercial or semi-commercial settings [3], [4]. These studies are important because they demonstrate that poultry behavior can be observed automatically in environments where bird overlap, motion, and scene complexity are common.

Other poultry-focused visual Artificial Intelligence studies have extended this idea beyond behavior observation toward health and condition assessment. For example, automated lesion detection in laying hens has been used within a precision livestock farming framework, showing that deep learning systems can be adapted to poultry-specific inspection tasks when the target condition is visually identifiable [5]. Real-time visual monitoring has also been applied to mortality-related detection in high-density broiler houses, further demonstrating that poultry Artificial Intelligence can address operationally meaningful events rather than only general observation tasks [16]. Related poultry Artificial Intelligence work also suggests that deployment efficiency matters alongside model accuracy when systems are intended for practical farm use rather than only laboratory evaluation [6].

The broader implication of this body of work is that poultry computer vision is already moving from proof-of-concept toward task-specific deployment. Still, the majority of existing applications focus on house monitoring, welfare indicators, disease-related cues, or bird condition assessment. Much less attention has been given to harvest-stage operations, where the purpose of monitoring is not only behavioral interpretation, but also count support, workflow visibility, and record verification assistance.

C. Dense Scenes, Counting, and Verification Tasks

One of the most persistent challenges in poultry computer vision is the dense visual arrangement of birds. Overlapping bodies, partial occlusion, irregular orientation, and crowd movement can reduce the reliability of standard object detection pipelines. This is especially relevant when the desired output is not merely bird presence, but a count or verification decision that may affect farm records or payment-related documentation. From a computer science perspective, this is the kind of setting where detector choice, annotation strategy, and model efficiency become inseparable from deployment value.

Recent research has addressed this issue more directly. Dense chicken counting work, such as PCCNet and DFCC-Net, shows that poultry-specific counting in crowded scenes

requires methods designed for high-overlap visual conditions rather than relying on simplified open-scene assumptions [7], [8]. Broiler-specific real-farm datasets such as PIO also strengthen this literature by providing more realistic visual benchmarks for evaluating detection systems under practical poultry farming conditions [9]. This line of work is particularly relevant to the present study because harvest-stage weighing and count verification can create local crowding, partial overlap, and irregular head positions that are visually different from birds spread across litter floors or feeding lanes.

Related work in automated broiler weighing further strengthens the case for treating monitoring as an operational task rather than a purely technical one. Automated precision weighing research has shown that video-derived features can support production-related measurements in ways that are directly meaningful to commercial operations [10]. That is important here because harvest-stage monitoring also revolves around a farm process with immediate management value. In both cases, the utility of the system depends on whether the output can be trusted within a real workflow, not only on whether it performs well on image-level benchmarks.

The literature on piling behavior detection also contributes an additional lesson: automated poultry monitoring becomes most valuable when it addresses a specific farm concern with clear action relevance [4]. In other words, good systems are not only accurate; they are also operationally interpretable. This principle is directly applicable to harvest-stage count verification, where the value of detection lies in supporting reliable verification during the process.

D. Oriented Object Detection and Practical Deployment Considerations

At the model level, this is where object-detection design becomes central. YOLO-based detectors are commonly adopted in recent poultry computer vision studies because they support efficient inference and practical field deployment [5], [6], [14]. For field deployment, that efficiency matters because count verification loses value if inference cannot keep pace with the harvest process.

However, standard axis-aligned boxes are not always ideal when objects are rotated, partially overlapping, or visually entangled. Oriented bounding box detection has therefore become an important technical option for cluttered scenes because it can represent object geometry more precisely than conventional horizontal boxes in such settings [17]. Although much of the OBB literature comes from other domains, the underlying problem of irregularly oriented, densely packed targets is highly relevant to poultry harvest imagery.

Practical deployment introduces an additional consideration. In poultry operations, the value of a model depends not only on image-level accuracy but also on whether detector size and runtime demands remain reasonable for real farm use. Existing deployment-oriented studies suggest that efficiency should be interpreted alongside model accuracy rather than treated as a completely separate concern [6]. For harvest-stage monitoring, detector choice and annotation strategy therefore need to be considered together with basic practical constraints.

E. Scenario-Specific Development in the Author's Prior Studies

The present work also builds on a sequence of scenario-specific studies rather than appearing in isolation. In prior work, the author compared standard axis-aligned bounding boxes and oriented bounding boxes for broiler detection in feeding and drinking scenes under dense commercial conditions [14]. That study highlighted an important practical point: the geometry used for annotation and detection can materially affect how well a model represents crowded poultry scenes.

A second study then examined camera placement and camera-to-subject distance, showing that deployment choices strongly influence the usefulness of poultry Artificial Intelligence in commercial houses [15]. Together, these studies suggest that poultry Artificial Intelligence should not be evaluated as a single universal solution. Instead, it should be assessed in relation to the exact farm scenario in which it will be used. Feeding and drinking scenes, general house monitoring, and harvest-stage weighing are not equivalent visual problems, even if they all involve broilers.

That scenario-specific view is important for farm audiences. Owners care whether a system works in their actual workflow. Veterinarians care whether the monitoring context is relevant to bird handling and welfare. Workers care whether the system reflects the real conditions in which they do their jobs. The present study extends the earlier line of work into harvest operations, where these practical questions become especially visible.

F. Research Gap

Taken together, the existing literature shows that poultry computer vision can support welfare observation, behavior analysis, lesion detection, dense-scene counting, and certain production-related measurements. It also shows that practical deployment, real-farm datasets, and task-specific design matter in real-world poultry systems [3], [5]–[10]. However, the literature remains thin in one specific area: harvest-stage monitoring for manual weighing and count verification.

This gap is significant because harvest is one of the few stages where welfare, labor, verification, and time pressure all converge. Existing work does not adequately address whether poultry computer vision can support manual harvesting workflows in farms that still depend on human handling rather than conveyor-based automation. It also does not sufficiently examine how object detection can support weighing and count verification in a way that is operationally useful for workers, owners, and veterinarians.

The present study addresses that gap by evaluating YOLO-based oriented bounding box detectors in a harvest-stage weighing and count verification scenario. The goal is to assess whether Artificial Intelligence can serve as a practical layer of operational support in a setting where better verification, clearer documentation, and shorter handling time may all matter. This gives the study both technical relevance and applied value within commercial poultry production.

III. METHODOLOGY

This study was designed as a scenario-specific applied computer vision evaluation for commercial poultry harvest operations. The objective was not only to measure image-level detection quality, but also to determine whether oriented bounding box (OBB) broiler-head detection could function as a practical verification layer during manual harvest weighing and record verification. In line with the framing of this paper, the methodology focuses on three linked questions: whether broiler heads can be localized reliably under harvest-stage visual conditions, whether those detections can support frame-level counting, and whether the resulting system can support event-level count verification in a farm setting.

A. System Overview

The proposed workflow was developed around a manual broiler harvest process in which birds are brought to the weighing area and visually observed near the point of record verification. A camera installed near the weighing station captures the hanging broilers while a computer vision detector performs automated detection of broiler heads. The detected heads are then treated as the visual basis for count verification support during the weighing process.

This workflow was intentionally designed as an assistive verification system rather than a fully autonomous replacement for workers. In practical terms, the system is intended to provide an additional counting and monitoring layer during a labor-intensive harvest stage, while remaining simple enough for practical use in real farm harvest settings. For this reason, the study emphasizes detector accuracy, operational visibility, and count-verification relevance alongside practical usability. The same setup may also support future time-and-motion analysis of harvest activities and weighing-station workflow efficiency.

B. Harvest-Stage Image Acquisition

The image acquisition setup was designed specifically for the harvest-stage weighing scenario addressed in this paper. Video and image data were collected from a commercial broiler harvest environment using a HIKVision E-HWIB camera with a 2.8 mm lens. The camera was positioned below the hanging broilers at an approximate working distance of 2.5 to 3 ft from the target area so that the broiler heads remained visible near the weighing point.

To improve image consistency during harvest operation, a custom rectangular enclosure was used around the camera system. The camera was mounted at the center of this enclosure, and two 5 W warm white LED lights were installed beside the camera to improve scene illumination. This lighting arrangement was intended to reduce shadowing and visibility loss during manual handling, while keeping the setup simple enough for practical farm use.

The capture system is based on a 2 megapixel camera stream. Frame dimensions, extraction settings, and the final dataset breakdown are reported in the experimental setup section. In total, 25,357 harvest-stage frames were extracted



Fig. 1. Field setup used for harvest-stage image acquisition, showing the illuminated enclosure positioned beneath suspended broilers at the weighing station.



Fig. 2. Sample frame from the installed camera view during manual broiler harvest, showing suspended broilers under field operating conditions at the weighing station.

during collection and screening. After filtering for annotation usability and task relevance, 3,935 raw images were retained and annotated for OBB broiler-head detection. Additional training-set expansion is handled separately through augmentation during model development rather than by treating transformed copies as newly collected field images.

C. Annotation Target and Labeling Strategy

The annotation target in this study is the broiler head only. Each identifiable broiler head is labeled using an oriented bounding box rather than a conventional axis-aligned box.



Fig. 3. Example annotated harvest-stage frame illustrating the broiler-head target under overlapping birds, partial occlusion, and irregular head orientations.

This design choice follows the visual conditions of the harvest environment, where full bodies are frequently overlapping, partially outside the effective field of view, suspended at irregular angles, or obscured by adjacent birds and handling movement. Under these conditions, the head serves as a more stable and operationally meaningful unit for count verification than the entire bird body.

The use of OBB labels was selected to better represent the irregular orientation of broiler heads during manual harvest handling. Compared with standard horizontal boxes, oriented boxes provide tighter geometric representation in crowded and rotated scenes, which is especially relevant when the downstream objective is reliable counting support rather than broad scene detection alone.

The labeling rule is intentionally narrow: only broiler heads are annotated. Annotation was performed offline using CVAT, and for consistency the full dataset was labeled by a single annotator. Detailed annotation rules for partially visible heads, severely blurred heads, and ambiguous cases are specified together with the dataset protocol in the experimental setup section. This is important because count verification performance depends not only on model quality but also on annotation consistency in dense scenes.

D. Modeling Approach

The detector family evaluated in this study is YOLO-based OBB detection. Exact architecture names, model scales, initialization details, and implementation identifiers are stated explicitly in the experimental setup section for reproducibility. At the methodological level, the system is treated as a YOLO-based oriented object detector intended for broiler-head localization under harvest-stage conditions.

The decision to use a YOLO-based detector follows the practical requirements of the study. The target system must support timely inference and maintain usable performance under dense, cluttered, and irregular visual conditions. These

constraints make detector design inseparable from practical usefulness in the present application.

E. Training and Data Preparation Strategy

Model training follows the same general hardware environment used in the author’s previous poultry computer vision studies. The training hardware specification, software environment, image size, training schedule, optimization settings, and model initialization details are reported explicitly in the experimental setup section to avoid ambiguity. These values are stated directly there rather than referenced indirectly from earlier work.

The study uses a standard supervised detection workflow with separate data partitions for model fitting, model selection, and final performance reporting. After frame extraction and dataset filtering, the usable annotated subset is organized using an approximately 70/15/15 train/validation/test split, and the evaluation partitioning protocol is defined explicitly in the experimental setup section. This decision matters because random frame-level splits may overestimate performance if visually similar frames from the same harvest sequence appear across multiple partitions.

Offline augmentation is applied only to the training partition so that validation and test performance remain based on held-out raw images. The augmentation pipeline is aligned with realistic farm-image variation, including changes in orientation, illumination, motion blur, scale, and partial occlusion. Any applied augmentations are therefore justified in relation to actual harvest conditions rather than generic detector training practice.

F. Detection-to-Count Verification Logic

This study is not intended to stop at object localization alone. The practical value of the system depends on whether broiler-head detections can support count verification during weighing and record verification. Accordingly, the methodology treats detection output as an intermediate result that must be converted into a usable counting signal.

At the image level, the detector outputs OBB head predictions subject to the usual confidence filtering and overlap suppression process. At the operational level, those detections are used to support verification of how many birds are visibly present in a short harvest-stage observation interval. The study therefore considers both image-level detection behavior and event-level count verification behavior.

At the frame level, each image is treated as an independent observation for detector evaluation and per-frame count comparison. For operational verification, however, the methodology does not rely on a single frame alone. Instead, short sequences associated with a weighing or verification event are analyzed as event windows so that temporary occlusion, motion blur, and frame-level detection instability do not dominate the final count estimate.

Within each event window, the Artificial Intelligence-assisted count is derived by aggregating valid frame-level detections across the sequence using a robust summary statistic. This event-level aggregation is intended to provide a more

stable verification signal than any single-frame estimate and is more closely aligned with the actual harvest workflow, where birds are observed over a short handling interval rather than at one isolated instant.

G. Evaluation Dimensions

The study is structured around three evaluation dimensions that correspond directly to the claims made in the opening sections of the paper.

First, image-level detection performance is measured to determine whether the model can localize broiler heads under harvest-stage visual conditions. This includes detector-quality measures such as precision, recall, and mean average precision under the defined implementation and partitioning setup.

Second, counting reliability is measured to determine whether the detections are operationally useful for verification. The metric set includes count error measures and agreement-oriented measures such as exact count match or near-match performance, because these are more interpretable in a poultry operations context than detector metrics alone.

Third, event-level count verification is assessed to determine whether short-window aggregation of frame-level detections can produce a stable verification signal that remains useful under temporary occlusion, motion blur, and handling-related instability.

H. Methodological Position of the Study

The methodological position of this work is intentionally applied. The paper does not claim a novel detector architecture. Instead, it evaluates whether a YOLO-based OBB detection pipeline can serve as a welfare-aware count verification layer for manual broiler harvest operations. This distinction is important. A weak version of the study would report detector accuracy alone. The stronger version, which this paper aims to present, evaluates the operational tradeoff among localization quality, frame-level counting behavior, and event-level count verification under a real poultry harvest scenario.

IV. EXPERIMENTAL SETUP

This section summarizes the dataset, annotation rules, model configurations, training procedure, event-level counting protocol, evaluation metrics, and computing environment used in the study.

A. Dataset Summary

The dataset was collected from commercial broiler harvest operations in Morong, Rizal and Baras, Rizal. Data collection covered three harvests across two collection sessions. Video was recorded at a source resolution of 1920×1080 , and frames were extracted at approximately 2 s intervals for dataset construction. In total, 25,357 harvest-stage frames were extracted during collection and screening. After filtering for annotation usability and task relevance, 3,935 raw images were retained and annotated for OBB broiler-head detection. The annotated subset contains 24,863 head instances, corresponding to an average of 6.32 labeled heads per annotated image.

TABLE I
DATASET SUMMARY

Item	Value
Collection sites	Morong, Rizal; Baras, Rizal
Harvests	3
Collection sessions	2
Source video resolution	1920 × 1080
Frame extraction interval	~2 s
Total extracted frames	25,357
Usable raw annotated images	3,935
Total OBB head annotations	24,863
Average labels per image	6.32
Filtering note	Retained after usability screening
Partitioning basis	Raw image grouped split (~70/15/15)

The raw annotated dataset was prepared for supervised detection training, validation, and testing, while augmentation was reserved for the training partition only. The filtering stage excluded frames that were unsuitable for reliable annotation or operational evaluation, such as frames with insufficient target visibility or limited usefulness for the broiler-head detection task. This conservative screening step was intended to preserve annotation reliability and task relevance under dense, cluttered harvest-stage conditions. The retained raw dataset was partitioned using an approximately 70/15/15 grouped split, corresponding to 2,739 training images, 584 validation images, and 612 test images. In this study, a raw image grouped split means that near-adjacent frames from the same short harvest sequence were assigned to only one partition, reducing leakage from visually similar images across training, validation, and test sets. Event-level verification analysis was based on short temporal windows derived from the harvest video sequences.

B. Annotation Protocol

Annotation was performed offline in CVAT using OBB labels, with broiler heads defined as the only annotation target. A head was annotated when the defining head structures were sufficiently visible to support reliable identification under harvest-stage conditions. Fully visible heads were labeled when the comb, eyes, beak, and wattles were visible. Partially visible heads were labeled when any two of these features remained visible and identifiable. Non-broiler objects, including workers, equipment, and other background elements, were ignored. Each visible head was labeled independently on a per-frame basis.

Although the harvest environment contained overlap, motion, and temporary occlusion, the labeling policy remained intentionally conservative and head-focused so that the annotation target aligned with the downstream count verification objective.

C. Model Configurations and Training Setup

The study evaluated YOLO-based OBB detectors trained from scratch on the harvest-stage dataset. Current project model identifiers include `yolo26n-obb.pt`,

TABLE II
ANNOTATION RULES

Visual case	Annotation rule
Fully visible head	Annotate when comb, eyes, beak, and wattles are visible
Partially visible but identifiable head	Annotate when any two of comb, eyes, beak, or wattles are visible
Non-broiler objects	Ignore
Visible head across frames	Annotate independently per frame

TABLE III
TRAINING SETUP

Item	Value
Detector family	YOLO-based OBB detection
Current project identifiers	<code>yolo26n-obb.pt</code> , <code>yolo26s-obb.pt</code> , <code>yolo26m-obb.pt</code>
Initialization	Random initialization; trained from scratch
Training hardware	Lenovo Legion 5 17ITH6
CPU	Intel Core i7-11800H
GPU	NVIDIA GeForce RTX 3050 Ti Laptop GPU
System memory	16.0 GiB
Software stack	Python, PyTorch, CUDA
Maximum epochs	300
Batch size	4
Early stopping patience	10
Augmentation focus	Rotation $\pm 15^\circ$, translation 5%, scale 0.90–1.10, brightness $\pm 20\%$, contrast $\pm 15\%$, gamma 0.85–1.15, motion blur 3–7 px

`yolo26s-obb.pt`, and `yolo26m-obb.pt`. In this paper, these filenames refer to the experiment-specific detector checkpoints associated with the nano, small, and medium OBB model configurations used in the study. The comparison results reported in Section V therefore correspond to the saved run identifiers rather than to separate benchmark families outside the present implementation. Although the checkpoint names follow the YOLO26-OBB naming convention, the models in this study were randomly initialized and trained only on the harvest-stage broiler-head dataset; pretrained external OBB weights were not used.

Training was conducted on a Lenovo Legion 5 17ITH6 laptop equipped with an 11th Gen Intel Core i7-11800H processor, 16.0 GiB RAM, and an NVIDIA GeForce RTX 3050 Ti Laptop GPU. Training runs were configured for up to 300 epochs with a batch size of 4 and patience of 10 epochs for early stopping. The training pipeline was implemented in Python using a YOLO-based OBB framework with PyTorch and CUDA support.

D. Augmentation Strategy

Augmentation was designed to reflect realistic harvest-stage variation while preserving the visual structure of the broiler-head detection task. The applied transformations were

TABLE IV
DATASET PARTITION SUMMARY

Partition	Raw images	Pre-training images	Share of retained set
Training	2,739	8,217	69.6%
Validation	584	584	14.8%
Test	612	612	15.6%
Total	3,935	9,413	100%

generated offline for the training partition only and were not applied to validation or test images. This design preserves a clean held-out evaluation while still increasing training diversity under conditions that may plausibly occur during harvest handling and camera-based observation, rather than broad synthetic perturbations with weak operational relevance. For the rotation component, each raw training image was retained in its original form and supplemented with two additional rotated variants, expanding the training set from 2,739 raw images to 8,217 images after offline rotation-based augmentation.

The augmentation pipeline included in-plane rotation of up to $\pm 15^\circ$, translation of up to 5%, scale variation from 0.90 to 1.10, brightness variation of $\pm 20\%$, contrast variation of $\pm 15\%$, gamma variation from 0.85 to 1.15, and motion blur with kernel sizes ranging from 3 to 7 pixels. These settings were intended to simulate realistic variation in suspended bird orientation, slight positional shift within the frame, modest scale change, illumination fluctuation, and motion-related image degradation during manual harvest operations.

E. Event-Level Count Verification Protocol

Detection quality was evaluated at the image level, but operational verification was also assessed over short temporal windows associated with harvest handling events. In this study, one event window was defined as an approximately 5 s sequence from the video stream near a weighing or verification moment. Within each event window, valid frames were those in which the suspended broilers remained visible within the intended working view of the camera.

For each valid frame, the frame-level Artificial Intelligence count was computed as the number of detected broiler heads after confidence filtering and overlap suppression. The event-level Artificial Intelligence-assisted count was then computed as the median frame-level count across the valid frames in the window. The reference event count was defined by manual review of the valid frames within the same event window, using visible broiler heads and frame-to-frame consistency as the basis for the final count label. This aggregation rule was selected to reduce the effect of transient motion blur, temporary occlusion, and short-lived frame-level detection instability during manual harvest handling.

F. Evaluation Metrics

Evaluation was structured along three dimensions: detection accuracy, frame-level counting reliability, and event-level count verification. Detection performance was quantified

TABLE V
EVALUATION DIMENSIONS AND METRICS

Dimension	Metrics
Detection accuracy	Precision, recall, mAP50, mAP50–95
Frame-level counting	MAE, RMSE, exact count match rate
Event-level verification	Event MAE, exact count match rate, ± 1 match rate

using precision, recall, mAP50, and mAP50–95. Counting performance was evaluated both at the frame level and at the event level. Frame-level counting emphasized per-image count comparison, whereas event-level counting emphasized the agreement between the aggregated Artificial Intelligence-assisted count and the reference count within each event window.

For counting reliability, the metric set includes mean absolute error (MAE), root mean squared error (RMSE), exact count match rate, and ± 1 count match rate where appropriate. Model scale was considered during interpretation because practical farm deployment generally favors compact detectors when accuracy differences are small.

V. RESULTS AND DISCUSSION

This section presents the quantitative and qualitative results, explains observed behavior, and discusses limitations.

A. Image-Level Detection Performance on the Validation Partition

The first stage of analysis focused on image-level oriented bounding box (OBB) detection performance on the held-out validation partition. At this point in the study, these results should be interpreted as model-selection evidence rather than as the final unbiased performance claim, because the separate test partition is reserved for later reporting. The purpose of the present comparison is therefore to determine which YOLO-based OBB configuration is most suitable for subsequent count-verification analysis under harvest-stage visual conditions.

For interpretive clarity, precision in this subsection refers to how often reported detections were correct, while recall refers to how often visible broiler heads were successfully detected. AP50 summarizes detection quality at a moderate overlap threshold, whereas AP50–95 provides a stricter measure across multiple thresholds and therefore better reflects localization consistency. In the context of harvest-stage count verification, higher recall reduces missed heads, higher precision reduces false detections, and higher AP50–95 indicates more stable OBB fitting under dense and irregular field conditions.

Three detector variants were evaluated: yolo26n-obb.pt, yolo26s-obb.pt, and yolo26m-obb.pt. These correspond to progressively larger model capacities, allowing the study to examine whether additional model complexity yields a meaningful improvement in broiler-head localization under dense, cluttered, and irregularly oriented harvest imagery. The

TABLE VI
VALIDATION DETECTION RESULTS

Model	Epoch	P	R	AP50	AP50-95
YOLO26n-OBB	84	0.9495	0.9541	0.9856	0.9266
YOLO26s-OBB	64	0.9570	0.9538	0.9865	0.9352
YOLO26m-OBB	39	0.9473	0.9573	0.9854	0.9287

comparison is especially relevant because practical farm deployment still favors compact detectors when performance differences are small.

Overall, all three models performed strongly on the validation partition. Precision and recall values remained consistently high across the best observed checkpoints, and all models produced mAP50 values close to 0.99. This indicates that the retained harvest-stage dataset, together with the offline augmentation strategy, was sufficient for all tested detectors to learn the broiler-head OBB detection task effectively. The primary question was therefore not whether the models could solve the task at all, but whether increasing model size materially improved localization quality under the stricter mAP50-95 criterion.

The `yolo26n-obb.pt` model achieved a best validation mAP50-95 of 0.9266 at epoch 84. At that checkpoint, the corresponding precision, recall, and mAP50 values were 0.9495, 0.9541, and 0.9856, respectively. Even the smallest detector in the comparison therefore demonstrated strong harvest-stage localization capability. This is important because the nano configuration represents the most compact model among the tested variants. Its performance suggests that accurate broiler-head OBB detection is achievable without immediately resorting to larger and more computationally expensive architectures.

The `yolo26s-obb.pt` model produced the strongest overall validation result, reaching a best mAP50-95 of 0.9352 at epoch 64. At that best checkpoint, precision was 0.9570, recall was 0.9538, and mAP50 was 0.9865. In practical terms, this result means that the small model provided the most favorable balance between localization quality and model scale in the present validation-stage comparison. Relative to the nano model, the increase in mAP50-95 was modest in absolute magnitude, but still meaningful because it reflects tighter and more consistent localization under a stricter evaluation threshold. This is the kind of improvement that matters when the downstream use case depends on reliable counting support rather than coarse object presence alone.

The `yolo26m-obb.pt` model remained highly competitive, with a best validation mAP50-95 of 0.9287 at epoch 39. At that checkpoint, precision was 0.9473, recall was 0.9573, and mAP50 was 0.9854. The medium model therefore maintained strong overall accuracy, but it did not surpass the small variant on the stricter localization metric. This outcome is notable because it shows that increasing model capacity beyond the small configuration did not produce a proportional gain in validation performance for the present harvest-stage task. In other words, the largest tested model did not clearly justify its additional complexity through a corresponding improvement in held-out validation accuracy.

The comparison also shows that the absolute differences

among models are relatively narrow. All three variants achieved high precision, high recall, and strong mAP50 values, while the separation in mAP50-95 was limited to a small but interpretable range. This pattern suggests that the task itself is already well captured by relatively compact OBB detectors once the training data, annotation target, and augmentation strategy are appropriately aligned with the harvest-stage scenario. The result is encouraging from an applied perspective because it implies that practical detector selection can prioritize compact model design without incurring a severe loss of image-level detection quality.

A second observation concerns convergence behavior. The best validation checkpoints appeared well before the nominal 300-epoch ceiling for all three runs: epoch 84 for nano, epoch 64 for small, and epoch 39 for medium. This suggests that the present dataset and training configuration allow the models to reach stable performance comparatively early. From an engineering standpoint, this is useful because it indicates that prolonged training beyond the plateau region may not be necessary to obtain competitive harvest-stage detection performance. From a methodological standpoint, it also strengthens confidence that the reported validation behavior reflects meaningful convergence rather than unstable late-epoch fluctuation.

Taken together, the validation-stage comparison supports three conclusions. First, YOLO-based OBB detectors are capable of localizing broiler heads reliably under the crowding, overlap, and irregular orientation patterns present in harvest-stage imagery. Second, the small model delivered the best overall validation performance among the tested configurations, making it the strongest current candidate for subsequent task-level analysis. Third, the medium model did not produce a clear validation advantage over the small model, while the nano model remained surprisingly competitive given its lighter computational profile. This ranking is operationally important because the study is not framed as detector novelty for its own sake; it is framed around practical count-verification value under farm-side constraints.

At the current stage of the study, the validation results therefore narrow the model-selection question rather than fully close it. The strongest provisional conclusion is that `yolo26s-obb.pt` offers the most attractive balance of accuracy and practicality for the harvest-stage detection problem, while `yolo26n-obb.pt` remains a viable compact alternative where lighter model design is preferred. Final conclusions about comparative model suitability, however, should still be confirmed against the reserved test partition and the later event-level counting analyses. For this reason, the present subsection should be read as the first half of the results story: a validation-grounded detector comparison that prepares the basis for the more operational evaluation steps that follow.

B. Held-Out Test Results for Detection and Count Verification

The final model comparison was performed on the reserved test partition, which contains 612 frames grouped into 216 short event windows. Unlike the validation results in the previous subsection, these test results are intended to support the main performance claims of the study because the partition

TABLE VII
HELD-OUT TEST DETECTION RESULTS

Model	Precision	Recall	mAP50	mAP50-95
YOLO26n-OBB	0.9435	0.9610	0.9675	0.9140
YOLO26s-OBB	0.9557	0.9507	0.9592	0.9156
YOLO26m-OBB	0.9479	0.9580	0.9684	0.9191

was not used for model selection during training. The test analysis jointly considered image-level detection quality, frame-level counting error, and event-level count verification so that detector quality could be interpreted in terms of operational usefulness rather than localization accuracy alone.

The test results show that all three models retained strong image-level detection performance on unseen harvest-stage data. Precision remained above 0.94 for all models, recall remained above 0.95 for the nano and medium variants and only slightly below that threshold for the small variant, while mAP50 remained above 0.95 across the board. These values indicate that the three detectors generalized well to the held-out test imagery and that the OBB formulation remained effective under dense, cluttered, and irregularly oriented harvest scenes.

In pure detection terms, the `yolo26m-obb.pt` model produced the strongest localization result on the test partition. It achieved the highest mAP50 of 0.9684 and the highest mAP50-95 of 0.9191, while maintaining a precision of 0.9479 and a recall of 0.9580. This confirms that the medium model remained highly capable as an object detector. However, the margin over the smaller models was narrow, particularly given the stronger downstream counting performance of the lighter variants. The result therefore reinforces an important practical point: modest gains in localization quality do not necessarily translate into the best count-verification behavior under operational farm conditions.

For frame-level counting, the `yolo26s-obb.pt` model provided the best overall performance. It achieved the lowest frame MAE of 0.4902 and the lowest frame RMSE of 0.8363, together with the highest exact frame match rate of 0.5980 and the highest ± 1 frame match rate of 0.9265. In practical terms, this means that the small model gave the most reliable per-frame broiler-head counts on the unseen test partition. Because manual verification during harvest may depend on stable frame-to-frame support rather than only on detector confidence, this result is operationally important and suggests that the small model provides the strongest general balance between detection quality and counting usefulness.

At the event level, the `yolo26n-obb.pt` model produced the lowest counting error. It achieved an event MAE of 0.4491 and an event RMSE of 0.7728, with an exact event match rate of 0.6204 and a ± 1 event match rate of 0.9352. Although the small model matched the nano model on the ± 1 event rate, the nano model performed better on the stricter error-based metrics and on exact event agreement. This indicates that the compact nano model remained especially effective once frame-level predictions were aggregated over short harvest windows using the event-level median rule.

The comparison also highlights that stronger detection mAP

did not automatically yield the strongest counting outcome. The medium model ranked first on mAP50 and mAP50-95, yet it produced the weakest frame-level and event-level count error among the three variants. Conversely, the small and nano models delivered better counting behavior despite slightly lower localization scores. This pattern suggests that for harvest-stage count verification, the most useful detector is not necessarily the one with the highest localization metric alone, but rather the one whose detection behavior remains most stable under the aggregation rules and operational conditions of the task.

Taken together, the held-out test results support a more nuanced model ranking than the validation-stage comparison alone. The small model is the strongest overall choice when the main objective is reliable frame-by-frame counting support, because it produced the best frame-level error profile while still maintaining very strong detection accuracy. The nano model is the most attractive option when event-level verification is prioritized and lighter deployment cost is desirable, since it achieved the lowest event-level error while remaining competitive on all image-level metrics. The medium model remains the strongest pure detector in terms of mAP, but its weaker counting behavior makes it less compelling as the default operational model for the present harvest verification task.

Because the event-level analysis was based on 216 short harvest windows from the retained dataset, the results should be interpreted as evidence of operational feasibility under the tested farm conditions rather than as a universal performance guarantee across all harvest environments.

C. Qualitative Failure Cases

Although the test results were strong overall, qualitative inspection showed several recurring failure modes that are consistent with the harvest-stage visual environment. The most common source of undercounting was heavy head occlusion, especially when one or more broiler heads were partially hidden behind another bird’s body or wings. In these cases, the visible evidence sometimes fell below the threshold required for stable detection, even when the event itself still contained the correct count information across nearby frames.

Motion blur was another important failure mode. During active handling, suspended broilers sometimes moved rapidly enough to soften fine head structures such as the beak, eyes, and comb. This reduced localization confidence and occasionally caused missed detections in single frames. Extreme overlap among nearby heads could also create ambiguity, sometimes leading the detector to merge adjacent heads into one prediction or, less frequently, to place overlapping detections in a crowded region.

Temporary obstruction by workers’ hands, weighing equipment, or other foreground objects also affected performance. These obstructions sometimes hid valid target heads and sometimes introduced visual clutter near the region of interest. In addition, difficult lighting conditions, including shadowed regions and uneven illumination across the birds, occasionally reduced confidence for partially visible heads. These effects did not dominate the dataset, but they help explain why even

TABLE VIII
HELD-OUT TEST COUNTING AND VERIFICATION RESULTS

Model	Frame MAE	Frame RMSE	Frame Exact	Frame ± 1	Event MAE	Event RMSE	Event Exact	Event ± 1
YOLO26n-OBB	0.5327	0.8819	0.5719	0.9134	0.4491	0.7728	0.6204	0.9352
YOLO26s-OBB	0.4902	0.8363	0.5980	0.9265	0.4861	0.8079	0.5880	0.9352
YOLO26m-OBB	0.5817	0.9411	0.5425	0.9020	0.5463	0.8767	0.5509	0.9167

the strongest detector did not achieve perfect frame-level or event-level agreement.

From an operational perspective, these failure modes support the study’s use of event-level aggregation rather than reliance on a single frame. Short-window median aggregation reduced the effect of transient blur, brief occlusion, and temporary detector instability, which is one reason the nano model remained particularly competitive in event-level verification despite not having the strongest frame-level metrics. The qualitative review therefore reinforces the central conclusion of the study: harvest verification performance depends not only on image-level localization quality, but also on how detection behavior interacts with real handling conditions and the downstream counting protocol.

D. Practical Recommendation for Deployment

The combined results suggest that model selection for harvest-stage verification should not be based on detection mAP alone. Although the `yolo26m-obb.pt` model achieved the strongest localization metrics on the held-out test partition, it did not produce the strongest counting performance. For practical deployment, the `yolo26s-obb.pt` model is the best balanced default choice because it delivered the strongest frame-level counting accuracy while still maintaining very strong image-level detection performance. The `yolo26n-obb.pt` model remains an attractive compact alternative when event-level verification and lighter deployment cost are prioritized, since it produced the lowest event-level counting error after short-window aggregation.

This recommendation is especially relevant because the intended farm-side use case is not limited to object detection as an isolated task. In a broader operational workflow, the proposed vision-based verification approach could be coupled with digital weighing devices, including Bluetooth-enabled scales, to support more streamlined count-and-weight recording and cross-checking during harvest. Such integration may reduce manual transcription steps, improve traceability of harvest records, and shorten handling time during verification. In turn, a more efficient verification workflow has the potential to improve operational consistency and support better animal welfare by reducing unnecessary delay during manual harvest weighing and count confirmation.

VI. CONCLUSION

This study evaluated YOLO-based oriented bounding box detection for broiler-head localization and count verification during manual poultry harvest weighing. Using harvest-stage imagery collected under practical farm conditions, the work compared nano, small, and medium detector variants not only

in terms of image-level detection accuracy, but also in terms of frame-level counting reliability and short-window event-level verification performance. The held-out test results showed that all three models achieved strong detection performance on unseen data, confirming that OBB-based localization is a viable approach for monitoring broiler heads in dense and irregular harvest scenes.

The results also showed that the strongest pure detector was not necessarily the strongest operational counting model. Although the medium model achieved the highest localization metrics, it did not provide the best counting behavior. The small model produced the strongest frame-level counting performance and is therefore the best balanced default choice for practical deployment, while the nano model produced the lowest event-level counting error and remains an attractive compact alternative when event-level verification is prioritized. This result demonstrates that model selection for poultry harvest verification should include operational count metrics, not detection metrics alone.

Overall, the study shows that computer vision can support count verification during manual broiler harvest operations and can provide useful assistance beyond conventional image-level detection reporting. In future deployment, the proposed approach could be integrated with digital weighing devices, including Bluetooth-enabled scales, to support more streamlined count-and-weight cross-checking, improved record traceability, and reduced manual handling delay during verification. Future work should extend evaluation across more farms, harvest conditions, and camera setups, and should further investigate workflow integration and practical near-real-time deployment under commercial operating conditions.

ACKNOWLEDGMENT

The author expresses sincere gratitude to Dra. Luisa E. Dandan for her invaluable guidance as a subject matter expert in poultry science. Her expertise in commercial broiler production, harvest-stage handling, poultry house management, and animal welfare helped ensure that the study remained grounded in real farm conditions and aligned with practical broiler harvest operations.

The author also gratefully acknowledges Ms. Genalin C. Salomon for her contribution to dataset preparation and annotation. Her careful work in image labeling and review of broiler-head annotations was important in building the harvest-stage oriented bounding box dataset used in this study.

Special thanks are extended to Mr. Amiel Jose Miguel C. Salomon for his technical support during field preparation and data collection. He assisted in the assembly of the camera and lighting box, as well as in the setup and testing of the image-capture system used for harvest-stage recording. His

support helped ensure stable and usable visual data under field conditions.

The author is likewise deeply grateful to Mr. Angelo Raya, the Raya family, and their farm staff for allowing data collection within their commercial poultry farm and for their generous support throughout the fieldwork. Their cooperation, hospitality, and the accommodations they provided during overnight data-collection activities made it possible to conduct the study under real harvest-stage farm conditions.

REFERENCES

- [1] S. Neethirajan, “Rethinking poultry welfare—integrating behavioral science and digital innovations for enhanced animal well-being,” *Poultry*, vol. 4, 2025.
- [2] N. Coimbra da Silva, I. d. A. Nääs, J. d. S. G. Barros, and D. J. d. Moura, “Modeling broiler discomfort under commercial housing: Seasonal trends and predictive insights for precision livestock farming,” *Poultry*, vol. 4, 2025.
- [3] S. Dahal, X. Yang, B. Paneru, A. Dhungana, and L. Chai, “Automatic detection of foraging hens in a cage-free environment with computer vision technology,” *Poultry*, vol. 4, 2025.
- [4] D. B. Jensen, M. Toscano, E. van der Heide, M. Grønving, and F. Hakansson, “Comparison of strategies for automatic video-based detection of piling behaviour in laying hens,” *Smart Agricultural Technology*, vol. 10, p. 100745, 2025.
- [5] T. Bergamasco, A. Ambrosi, V. Tregnaghi, R. Urbani, G. Nalesso, F. Menegon, A. Trocino, M. Pravato, F. Bordignon, S. Sparesato, G. Manca, and G. Di Martino, “Precision livestock farming: Yolov12-based automated detection of keel bone lesions in laying hens,” *Poultry*, vol. 4, 2025.
- [6] Q. Tong, J. Wang, and W. Yang, “Edge ai-enabled chicken health detection based on enhanced fcos-lite and knowledge distillation,” *arXiv preprint arXiv:2407.09562*, 2024.
- [7] Y. Guo, Z. Wu, Z. Su, J. Zhao, and X. Li, “Pccnet: A point supervised dense chickens flock counting network,” *Smart Agricultural Technology*, vol. 10, p. 100795, 2025.
- [8] J. Lv, J. Wang, C. Peng, and Q. Huang, “Dfccnet: A dense flock of chickens counting network based on density map regression,” *Animals*, vol. 13, no. 23, p. 3729, 2023.
- [9] K. Boniche, E. Cruz, J. C. Rangel, M. Hidalgo-Rodríguez, and F. Gómez-Donoso, “Pio, a large-scale dataset for broiler chicken detection under real poultry farming conditions,” *Scientific Data*, 2026.
- [10] M. Campbell, P. Miller, K. Díaz-Chito, S. Irvine, M. Baxter, J. Martínez del Rincón, X. Hong, N. McLaughlin, T. Arumugam, and N. O’Connell, “Automated precision weighing: Leveraging 2d video feature analysis and machine learning for live body weight estimation of broiler chickens,” *Smart Agricultural Technology*, vol. 10, p. 100793, 2025.
- [11] M. A. Hettmannsperger and I. Ruhnke, “Manual catching and transportation of poultry with a focus on chickens and european practices,” *Poultry*, vol. 5, 2026.
- [12] K. E. Kittelsen, E. G. Granquist, A. L. Aunsmo, R. O. Moe, and E. Tolo, “An evaluation of two different broiler catching methods,” *Animals*, vol. 8, no. 9, p. 141, 2018.
- [13] I. Wolff, S. Klein, E. Rauch, M. Erhard, J. Mönch, and S. Härtle, “Harvesting-induced stress in broilers: Comparison of a manual and a mechanical harvesting method under field conditions,” *Applied Animal Behaviour Science*, vol. 221, p. 104877, 2019.
- [14] A. R. Salomon, “Machine learning in commercial poultry farming: Comparing standard and oriented bounding boxes in yolov11 for broiler detection,” *TechRxiv*, Apr. 2025.
- [15] —, “Machine learning in commercial poultry farming: Optimizing camera placement to bridge ai performance and monitoring in high-density broiler stocking,” *TechRxiv*, Apr. 2025.
- [16] L. Yang and J. Song, “Real-time detection of dead broilers in high-density poultry houses using high-resolution video and semantic segmentation,” *Poultry Science*, vol. 105, p. 106969, 2026.
- [17] K. Wang, Z. Wang, Z. Li, A. Su, X. Teng, M. Liu, and Q. Yu, “Oriented object detection in optical remote sensing images using deep learning: A survey,” *arXiv preprint arXiv:2302.10473*, 2024.