SMATIKA : STIKI Informatika Jurnal
Accredited Sinta 4 Ministry of Higher Education, Science and Technology Republic of Indonesia SK No. 79/E/KPT/2023
E-ISSN: 2580-6939 || P-ISSN: 2087-0256

YOLOv11-Based AIoT System for Automated Size Detection and Counting of G0 Seed Potatoes Using MQTT Protocol

Alvandi Fredik Sembiring [1], Indah Permatasari [2*], Prasetyo Yuliantoro [3]

[1,2,3] *Telkom University, Faculty of Electrical Engineering, Telecommunication Engineering, Jl. DI Panjaitan No. 128, Karangreja, Purwokerto Kidul, Purwokerto Sel. District, Banyumas Regency, Central Java 5314, Indonesia*

Keyword

AIoT; G0 Seed Potato; Object Detection; YOLOv11

Corresponding Author:
indahpermata@telkomuniversity.ac.id

Abstract

The increasing demand for G0 seed potatoes in Indonesia, reaching approximately 143,740 tons in 2021 while only 8.6% of the demand could be supplied, highlights the need for more efficient production and monitoring systems at the early stage of the potato seed supply chain. Current manual sorting and counting processes are labor-intensive, time-consuming, and prone to human error, creating a need for an automated and reliable monitoring solution. This study develops and implements an Artificial Intelligence of Things (AIoT) system based on the YOLOv11n object detection algorithm for real-time detection, size classification, and counting of G0 seed potatoes. The proposed system integrates a conveyor belt, a Logitech C922 Pro USB webcam for image acquisition, and a laptop as the edge computing unit running the YOLOv11n model. Detection results are transmitted through the MQTT protocol to a Node-RED dashboard for real-time remote monitoring. Unlike conventional approaches, the system combines a lightweight YOLOv11n model with MQTT communication to support simultaneous multi-category size classification and synchronized dashboard visualization. Detected potatoes are classified into three size categories (small, medium, and large) based on calibrated bounding-box pixel areas validated with potato farmers. The model was trained and evaluated using four epoch configurations (25, 50, 75, and 100 epochs) with Precision, Recall, F1-score, mAP@0.5, and mAP@0.5–0.95 as evaluation metrics. The 100-epoch model achieved the best performance, with precision approaching 1.00, recall of approximately 0.98, mAP@0.5 of 0.986, and mAP@0.5–0.95 of 0.96. Validation confirmed that calibrated geometric measurements matched the physical potato dimensions, while dashboard data were fully consistent with edge-computing outputs. These findings demonstrate that the proposed YOLOv11n-based AIoT system provides accurate, reliable, and real-time monitoring of G0 potato production, offering a practical solution to improve operational efficiency and data accuracy in Indonesia's potato seed supply chain.

1. Introduction

In real agricultural environments, monitoring G0 potato production presents significant practical challenges, including limited physical access in enclosed screenhouses or aeroponic chambers, the difficulty of manually classifying small and morphologically variable tubers consistently, and the absence of standardized automated tools that results in fragmented, non-real-time data recording across production facilities. These challenges are deeply rooted in the strategic importance of potato as a food commodity and the structural weaknesses of Indonesia's seed supply chain, which together create an urgent demand for more efficient and technology-driven production management.The Potato (*Solanum tuberosum* L.) is one of the most strategically important food commodities in Indonesia, serving as a staple food source and a key ingredient in the food processing industry. Global potato production consistently ranks fourth among major food crops after wheat, rice, and corn, with total world production exceeding 370 million tons per year [1]. Potato demand in Indonesia reaches approximately 6.2 million tons per year [2].However, the national potato seed supply chain faces critical challenges, particularly at the root generation level. G0 potato seeds, which are zero-generation seeds produced through tissue culture under sterile and controlled conditions, serve as the primary source for subsequent seed generations (G1–G4). Studies show that demand for potato seeds in Indonesia reached approximately 143,740 tons in 2021, while the available supply covered only approximately 8.6%, or around 12,361 tons [3]. Therefore, more efficient seed management is needed at every stage of the production chain. G0 seed potatoes are characterized by their small size (typically 2–10 grams), smooth skin surface, high genetic purity, and freedom from systemic viral infections such as PVX, PVY, and *Rhizoctonia solani*, with aeroponic production systems reportedly yielding 25–30% higher efficiency compared to conventional methods [4]. Despite their critical role in the seed supply chain, the sorting and counting of G0 potatoes remain predominantly manual processes, which are time-consuming, labor-intensive, and prone to human error. Several studies have emphasized that increased efficiency at the G0 stage can have a cascading positive impact on the overall productivity of the seed potato production chain [5].

The rapid advancement of digital technologies, particularly Artificial Intelligence (AI) and the Internet of Things (IoT), has opened up new opportunities for precision agriculture and smart farming applications. IoT-based systems are increasingly being implemented in agricultural environments to monitor environmental parameters such as soil temperature and moisture in real time [6]. Furthermore, studies have demonstrated the implementation of potato cultivation automation using aeroponic methods integrated with IoT technology [7]. The convergence of AI and IoT into what is now termed Artificial Intelligence of Things (AIoT) represents a highly promising paradigm for agricultural automation, combining the sensing and connectivity capabilities of IoT with the intelligent data processing capabilities of AI [8]. Among existing AI-based object detection frameworks, the You Only Look Once (YOLO) algorithm stands out for its real-time processing capability, high detection accuracy, and computational efficiency, as it processes the entire image in a single pass through a neural network, enabling simultaneous object detection and classification at high speed [9],[10]. A further study successfully applied YOLO to agricultural object detection with high accuracy, achieving a mean Average Precision (mAP) of approximately 91%, demonstrating the algorithm's adaptability to local agricultural contexts [11].

Several previous studies have explored the application of YOLO and other deep learning models for agricultural object detection and counting. A modified SCG-YOLOv8n framework has been developed for automated potato counting, achieving an R^2 value of 0.96 while supporting efficient mobile deployment. However, the system was limited to counting tasks and lacked IoT integration and multi-category size classification, preventing remote monitoring and direct grading functionality [12]. The effectiveness of YOLO-based object detection has also been demonstrated in real-time Android applications, achieving an accuracy of 0.98 [13]. In addition, a corn kernel counting system combining YOLOv8 with the ByteTrack tracking algorithm achieved a counting error of less than 3% under complex overlapping conditions, although it was designed solely for static batch counting without real-time IoT data transmission [14]. Another study introduced YOLOv8-HD for high-density wheat kernel detection and counting, achieving a mAP@0.5 of 77.6% with an inference time of only 2.86 ms per frame; however, the approach neither supported size classification nor incorporated remote monitoring capabilities [15]. Furthermore, CNN-based transfer learning combined with explainable AI has achieved validation and testing accuracies of 97% and 98%, respectively, for potato leaf disease classification, demonstrating the strong potential of deep learning for potato-related agricultural applications [16].

The integration of artificial intelligence with IoT technologies has also received increasing attention. A real-time crowd counting system combining YOLOv8 NAS with IoT sensors achieved an mAP of 95.1% while maintaining a response time of less than two seconds [17]. Similarly, an IoT-based automated fish seed counting prototype reported a maximum counting error of only 3.65% [18]. In addition, a comprehensive AIoT framework has been proposed for real-time crop yield prediction to support agricultural decision-making [8]. These studies demonstrate that integrating deep learning with IoT technologies enhances automation, monitoring capability, and decision support. Nevertheless, existing studies have not simultaneously addressed real-time potato size classification, object counting, and IoT-based remote monitoring within a single integrated system, which constitutes the primary contribution of the present study.

Despite these advances, significant research gaps remain. Most existing studies focus either on object detection and counting without IoT integration, or on IoT-based monitoring without AI-driven visual classification. Studies specifically addressing the automatic detection, size classification, and real-time counting of G0 seed potatoes integrated with an IoT monitoring dashboard remain scarce. The unique morphological characteristics of G0 potatoes, including their small and irregular shape, variable size, and visually similar appearance across size categories, present considerable challenges for automated visual inspection systems that have not been sufficiently addressed in the existing literature. Furthermore, the critical importance of accurate G0 potato inventory data for national seed supply chain management further motivates the development of reliable, automated, and remotely accessible monitoring systems. To address this gap, this study proposes and implements a novel AIoT system that integrates the YOLOv11n object detection model with the MQTT-based IoT communication protocol for the automatic detection and size-based counting of G0 seed potatoes. The system utilizes a conveyor belt mechanism equipped with a USB webcam to capture real-time images of G0 potatoes as they pass through the detection zone.

2. Research Method

2.1 System Architecture

This research was conducted through a series of systematic stages, as illustrated in the research flow diagram.

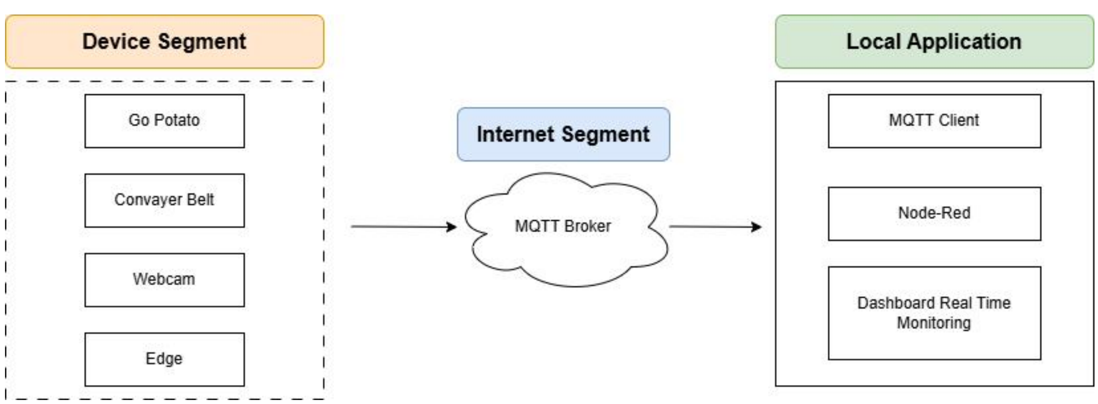

Figure 1. System Architecture

The overall system architecture of the proposed AIoT system consists of three main segments, namely the Device Segment, the Internet Segment, and the Local Application Segment, as shown in Figure 1. In the Device Segment, a G0 potato is placed on a conveyor belt and moves continuously toward the image capture area. A USB webcam mounted on the conveyor captures real-time images of the potato as it passes through the observation zone. The laptop connected to the webcam serves as the main processing unit, running the

YOLOv11n model for object detection, size classification, and potato counting. In the Internet Segment, detection data is transmitted via Wi-Fi using the MQTT protocol and published to the MQTT broker at broker.emqx.io. In the Local Application Segment, a second laptop acts as an MQTT subscriber that receives the published data and visualizes it through the Node-RED Dashboard, which can be accessed via a web browser for real-time monitoring purposes.

2.2 System Design and Implementation Setup

The hardware design consists of a motorized conveyor belt as a mechanical transport mechanism for the G0 potatoes, a Logitech C922 Pro Stream webcam mounted on a fixed bracket directly above the conveyor belt at a perpendicular angle to ensure consistent image capture geometry, and a laptop serving as an edge computing unit for real-time AI inference. The camera is connected to the laptop via a USB 3.0 interface and captures video frames at a native resolution that is subsequently resized to 640×640 pixels for YOLOv11n model inference. The software design was developed using Python as the primary programming language and structured into four functional layers: an image acquisition layer that leverages OpenCV for real-time capture of webcam video streams; an AI inference layer that implements the YOLOv11n model via the Ultralytics library with a confidence threshold of 0.25; an object tracking and counting layer that utilizes a custom Centroid Tracker algorithm with a maximum loss tolerance of 15 frames and a maximum tracking distance of 80 pixels to assign unique object IDs and detect left-to-right crossover events for accurate counting; and an IoT communication layer that employs the Paho-MQTT library to publish JSON-formatted detection payloads to the MQTT broker at broker.emqx.io on port 1883 under two dedicated topics, namely potato/count for individual detection events and potato/stats for aggregate statistics published every 5 seconds, with Node-RED configured on the subscriber side to receive and visualize the data via a real-time web-based monitoring dashboard.

2.3 Dataset Collection and Labeling

The dataset used in this study was manually collected using a camera to capture images of G0 potatoes under controlled laboratory conditions. A total of 100 images of G0 potatoes were collected for the dataset. Each image was captured individually to ensure sufficient variation in potato size, orientation, and lighting conditions. Following image collection, each image was annotated using Label Studio, an open-source data labeling tool, in which a bounding box was drawn around each potato and labeled with the class name "potato." The dataset was subsequently divided into two subsets, with 70 images allocated for training and 30 images allocated for validation, following a 70:30 split ratio to ensure sufficient data for model learning and performance evaluation. The utilization of a small dataset consisting of 100 highly curated images raises logical constraints regarding broader model generalization. However, the YOLOv11n architecture overcomes this limitation through two primary architectural advantages. First, the model leverages robust pre-trained weights from the COCO dataset via transfer learning, Second, this architecture can dynamically predict the center and edge coordinates of the potato. This makes guessing the position of the bounding box around the potato much more accurate and precise.

2.4 Potato Size Classification Criteria

The size classification of detected G0 potatoes was performed based on the pixel area of the bounding box generated by the YOLOv11n model. The pixel area was calculated by multiplying the width and height of the bounding box in pixels. Three size categories were defined as follows: potatoes were classified as Small (S) if their bounding box area was less than or equal to 45,000 pixels; potatoes were classified as Medium (M) if their bounding box area was greater than 45,000 pixels and up to 58,800 pixels; and potatoes were classified as Large (L) if their bounding box area exceeded 58,800 pixels. These pixel threshold values were determined through a validation process in which the pixel measurements of detected potatoes were cross-checked and confirmed by potato farmers to ensure alignment with the actual physical size categories.

2.5 YOLOv11n Model Training and Validation

YOLOv11 is the latest iteration of the widely used YOLO real-time object detection architecture. The YOLOv11 architecture consists of three main components: a Darknet-53 and CSPLayer-based Backbone for efficient feature extraction, a Neck with an FPN+PAN structure for multi-scale feature fusion, and a Head module that employs an anchor-free mechanism to optimize object classification and localization with greater accuracy and computational efficiency [19],[20]. This study employs the YOLOv11n model, which is the nano variant of the YOLOv11 architecture developed by Ultralytics, owing to its lightweight design and high inference speed, making it suitable for real-time detection on standard laptop hardware. The model was trained using Google Colaboratory with GPU acceleration to reduce training time. Training was performed under four different epoch configurations, namely 25, 50, 75, and 100 epochs, to systematically analyze the effect of training duration on model performance and to identify the optimal training configuration. All training configurations used an input image size of 640×640 pixels and a confidence threshold of 0.25 during inference. The training process involved iterative forward and backward propagation through dataset batches, with model weights updated after each batch based on the calculated loss value.

Model performance was evaluated using the following metrics derived from the confusion matrix framework. Precision measures the proportion of correct positive detections out of all positive predictions made by the model. Recall measures the proportion of true positive objects correctly identified by the model. The F1-Score represents the harmonic mean of Precision and Recall, providing a balanced measure of overall detection performance. Mean Average Precision at an IoU threshold of 0.5 (mAP@0.5) evaluates the average precision of bounding box predictions at a fixed Intersection over Union (IoU) threshold of 0.5. Mean Average Precision across IoU thresholds ranging from 0.5 to 0.95 (mAP@0.5–0.95) provides a more rigorous evaluation of localization accuracy across varying IoU thresholds. Training and validation loss values, including box loss, classification loss, and distribution focal loss, were also monitored throughout the training process to assess model convergence and detect potential overfitting

$$Precision = \frac{TP}{TP + FP} x100\% \quad (1)$$

$$Recall = \frac{TP}{TP + FN} x100\% \quad (2)$$

$$F1 - Score = \frac{2xPrecisionxRecall}{Precision + Recall} x100\% \quad (3)$$

$$mAP = \frac{1}{N} \sum_{i}^{N} AP_i \quad (4)$$

Where TP is True Positive, FP is False Positive, FN is False Negative, N is the number of classes, and AP is the Average Precision for each class.

2.6 IoT Integration and MQTT Communication

The IoT integration component of this system is implemented using the MQTT (Message Queuing Telemetry Transport) protocol, which is a lightweight publish-subscribe messaging protocol well-suited for IoT applications owing to its low bandwidth consumption and reliable message delivery [21]. After detecting and classifying G0-grade potatoes as they cross the centerline of the camera frame from left to right, the system immediately publishes a JSON-formatted data payload to the MQTT broker at broker.emqx.io over port 1883 using Quality of Service (QoS) level 0. The published payload includes the following fields: a timestamp indicating when the detection occurred, the total number of potatoes detected, the number per size category (small, medium, and large), an object ID as a unique identifier for each detected potato, a class name, a confidence score, bounding box coordinates, pixel area, label and size code, frame index, and movement direction. A separate MQTT topic is used for aggregate statistics, which are published immediately after each new detection event and periodically at regular 5-second intervals to ensure dashboard continuity. A Node-RED dashboard, running on the customer's laptop, receives these MQTT messages and visualizes them in real-time through graphical components including charts, gauges, and text displays.

Regarding communication security and message reliability, the current implementation configured the MQTT client using the Paho-MQTT library with MQTT_USERNAME and MQTT_PASSWORD set to None, indicating that no authentication, TLS encryption, or access control mechanism was applied, and message delivery was set to MQTT_QOS = 0 ("at most once"), which provides no acknowledgment guarantee between the publisher, broker, and subscriber, meaning that message loss may occur under unstable network conditions without any automatic retransmission. Acknowledge that communication security, guaranteed message delivery, and end-to-end network latency were not experimentally evaluated in this study and represent limitations of the current implementation that should be addressed in future work through the adoption of TLS encryption, client authentication, and higher QoS levels.

3. Result and Discussions

3.1 System Implementation Result

The results of the proposed AIoT system design for automatic detection and calculation of G0 potato size have been successfully implemented.

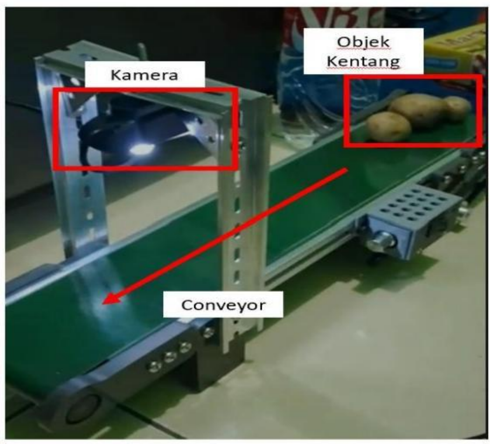

Figure 2. Hardware Design Results

Potatoes are positioned on a conveyor belt that moves continuously toward the image acquisition zone. A camera is mounted above the conveyor to capture real-time images of the potatoes. The camera is connected to a laptop via a USB interface to enable real-time data processing. The visualization of the potato size classification results is presented in the figure below.

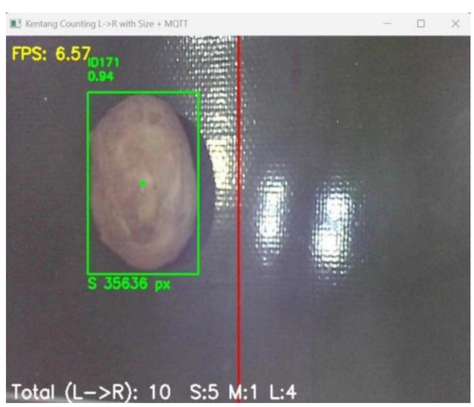

Figure 3. G0 Potato Detector and Counter Visual Display

The real-time visual output successfully validates the multi-category size grading and automated tracking logic as individual G0 potatoes pass through the detection zone. This seamless visualization confirms that the hardware-software integration functions reliably as a unified AIoT pipeline. Consequently, these deployment results demonstrate the technical feasibility and robust performance of embedding YOLOv11n-based computer vision [22] within a lightweight MQTT-based IoT infrastructure [23] designed specifically for agricultural automation purposes.

3.2 YOLOv11n Model Performance Evaluation and Discussion

The YOLOv11n model was trained and evaluated under four distinct epoch configurations (25, 50, 75, and 100 epochs) to systematically analyze the evolutionary trajectory of the network learning process and determine the optimal duration for the single-class detection task. To reduce structural redundancy and enhance the manuscript's scientific focus, the definitive quantitative metrics derived from the validation confusion matrices across all evaluated configurations are consolidated into Table 1.

Table 1. Summary of YOLOv11n Performance Metrics Across Epoch Configurations

Metric	Epoch 25	Epoch 50	Epoch 75	Epoch 100
Precision	1.00	1.00	0.9994	0.9995
Recall	0.98	0.99	0.98	0.98
F1-Score (max)	0.99	0.99	0.99	0.99
mAP@0.5	0.985	0.985 –0.99	0.985	0.986
mAP@0.5–0.95	0.91	0.93	0.94	**0.96**
val/box_loss	~0.36	~0.35	~0.29–0.30	~0.24–0.26
val/cls_loss	~0.50	~0.33	~0.20	~0.15–0.18
val/dfl_loss	~0.80	~0.80	~0.78	~0.77

The training trends demonstrate that even at the earliest configuration (epoch 25), the model exhibited strong initial convergence, achieving an mAP@0.5 of 0.985. This high initial performance indicates that the lightweight network quickly adapted to the high-contrast laboratory image boundaries. As training extended to 100 epochs, the model attained its peak generalization capability, where the mAP@0.5–0.95 progressively increased from 0.91 to an optimal 0.96. This total improvement of 5.5 percentage points is highly significant for the size classification task, as more precisely localized bounding boxes directly yield more accurate pixel area measurements, resulting in flawless physical size category assignments. The complete training and validation metric curves for the optimal 100-epoch configuration are illustrated in Figure 4.

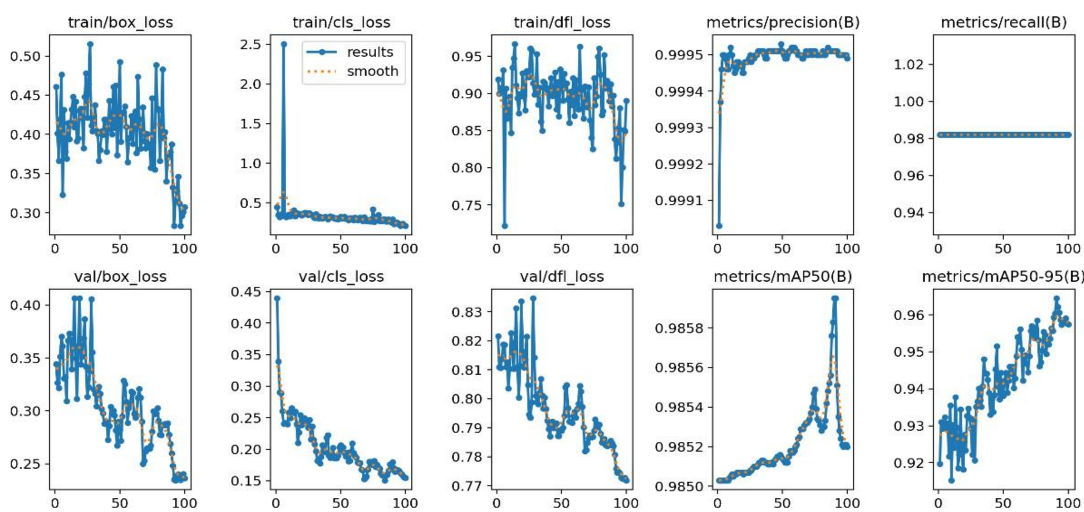

Figure 4. Training and validation metric curves at epoch 100

Overall, the results presented in Table 1 confirm three key observations. First, the precision metric remained consistently high across all epoch configurations, ranging from 1.00 at epochs 25 and 50 to 0.9994 and 0.9995 at epochs 75 and 100 respectively, confirming that the model produces very few false positive detections regardless of training duration, which is a critical requirement in agricultural sorting applications where incorrect size classifications could compromise inventory data integrity. Second, the recall value of 0.98–0.99 across all configurations confirms that the model successfully detected nearly all G0 potatoes present in each frame with minimal missed detections, ensuring reliable counting performance throughout all tested training configurations. Third, At epoch 100 mAP@0.5 of 0.986, and mAP@0.5–0.95 of 0.96, accompanied by the lowest validation loss values among all configurations, namely val/box_loss ≈ 0.24–0.26, val/cls_loss ≈ 0.15–0.18, and val/dfl_loss ≈ 0.77. The progressive improvement in mAP@0.5–0.95 from 0.91 at epoch 25 to 0.96 at epoch 100, representing a total gain of 5.5 percentage points, is particularly significant as it reflects the model's increasing capacity to produce precisely localized bounding boxes across multiple IoU thresholds, directly contributing to more accurate pixel area measurements and consequently more reliable size category assignments. the epoch 100 configuration is identified as the optimal training setting, offering the best combination of detection accuracy, bounding box localization precision, classification loss stability, and generalization capability, and is therefore selected as the recommended model configuration for deployment in the proposed AIoT system for G0 potato production monitoring.

The effective performance of YOLOv11n under the relatively small dataset condition used in this study can be attributed to several critical architectural and task-specific factors. From an architectural perspective, YOLOv11n employs compact C3k2 bottleneck blocks and a lightweight attention mechanism (C2PSA) in its backbone, which reduce the total number of trainable parameters compared to its predecessors while preserving strong feature extraction capability. This parameter efficiency is particularly advantageous when training data is limited, as models with fewer parameters are inherently less prone to overfitting and can converge to generalizable solutions with smaller training sets. A systematic benchmarking evaluation comparing the proposed system against recent object detection methodologies is presented in Table 2.

Table 2. Systematic Benchmarking with Prior Agricultural Detection Frameworks

Architecture	Multi-Category Size Grading	IoT Remote Pipeline Integration	Primary Reported Metric
Modified SCG-YOLOv8n Potato Counting [12]	No	No	$R^2 = 0.96$
Maize Kernel Batch YOLOv8 + ByteTrack [14]	No	No	$R^2 = 0.98$
Wheat Seed Detection and Counting YOLOv8-HD [15]	No	No	mAP@0.5 = 77.6%
Detection and Counting of G0 Seed Potatoes YOLOv11n	Yes	Yes	mAP@0.986

Based on the systematic benchmarking presented in Table 2, the proposed YOLOv11n system establishes a distinct operational advantage and analytical positioning compared to recent deep learning frameworks deployed in agricultural object tracking. Prior studies have focused primarily on isolated, localized vision tasks without network telemetry or structural classification. Specifically, the modified SCG-YOLOv8n model developed for potato counting and the YOLOv8 + ByteTrack framework optimized for maize kernel batch tracking achieve high centralized calculation benchmarks, reporting coefficient of determination R^2 values of 0.96 and 0.98, respectively. Similarly, the YOLOv8-HD framework tailored for wheat seed detection reports a mean Average Precision (mAP@0.5) of 77.6%. However, all three of these benchmarked methodologies function strictly as standalone, offline counting tools that completely lack multi-category size grading mechanisms and real-time IoT network pipeline integration. By embedding YOLOv11n architecture alongside a centroid tracking pipeline, the proposed system not only achieves a superior localization performance with a mAP@0.5 of 0.986, but it is also the framework that provides built-in multi-category size grading functionality and full IoT remote synchronization via the MQTT protocol.

3.3 Geometric Size Measurement and Calibration Validation

In real agricultural environments, several practical challenges must be acknowledged as potential sources of performance degradation in the pixel-area-based size classification approach, including varying illumination intensities and shadows, dust or moisture accumulation on the camera lens, overlapping potatoes on the conveyor belt, conveyor speed variations, and broader environmental variability in greenhouse or field operations such as temperature fluctuations and seasonal variations in potato morphology; therefore, to rigorously evaluate the scientific accuracy of the visual sensing system under the current controlled conditions, validation was performed by cross-referencing the pixel-area measurements generated by the YOLOv11n model against the actual physical dimensions of the G0 seed potatoes, where rather than treating the system as a generic black-box counter, this validation focused on the calibration accuracy of the non-contact bounding box tracking logic, with the actual classification of small, medium, and large categories manually verified by potato farmers based on physical measurements utilized as the definitive ground truth to test the system's threshold resilience.

Table 3. Bounding Box Pixel-Area Calibration and Size Classification Results

No	Actual Potato Dimensions (Ground Truth)	YOLOv11n Pixel-Area Prediction (Average Pixels)	Assigned Visual Category	Dashboard Sync Status
1.	Small (Physical Diameter ≤2.5 cm)	35,636 px	Small (S)	Synchronized
2.	Medium (Physical Diameter 2.6–3.5 cm)	49,120 px	Medium (M)	Synchronized
3.	Large (Physical Diameter ≥3.5 cm)	61,450 px	Large (L)	Synchronized

The measurement metrics in Table 3 prove that the calibrated pixel thresholds—Small (≤45,000 pixels), Medium (45,001–58,800 pixels), and Large (≥ 58,800 pixels)—maintain a robust mathematical correlation with the physical geometric boundaries of the crops. Even when confronting natural morphological irregularities, such as asymmetric surfaces or early-stage sprouting shoots, the anchor-free localization mechanism extracted the outer edge boundaries without experiencing spatial distortion.Furthermore, because the fixed perpendicular camera geometry ensures a constant distance between the lens and the conveyor belt surface, the calculated pixel area remained invariant across multiple runs. This stable geometric calibration guarantees that every data packet pushed to the remote Node-RED monitoring dashboard via the MQTT protocol represents an accurate physical sorting metric rather than an uncalibrated estimation, directly supporting precision inventory management for high-purity G0 seed supply chains. Although the consistent alignment between training and validation loss curves across all epoch configurations confirmed the absence of overfitting in this study, the relatively small dataset collected under controlled conditions limits the model's exposure to the full range of morphological and environmental variability encountered in real agricultural settings, and additional validation using larger and more diverse datasets sourced from multiple production facilities is still necessary to more rigorously assess the generalization capability of the proposed YOLOv11n-based system beyond the controlled experimental scope of this study.

When situated within the broader trajectory of YOLO-based agricultural object detection research, the performance achieved by the proposed YOLOv11n model corroborates and extends several findings reported in prior studies. The near-perfect precision and high recall obtained in this study are consistent with the high accuracy levels reported for real-time YOLO-based detection in other application domains, such as the 0.98

accuracy achieved in an Android-based currency detection system [13] and the strong validation performance reported for CNN-based potato leaf disease classification [16], suggesting that YOLO-family architectures, including its lightweight YOLOv11n variant, generalize well to small-object and fine-grained agricultural classification tasks even when trained on comparatively limited datasets. This pattern reinforces earlier observations that YOLO's single-pass detection mechanism enables simultaneous high-speed and high-accuracy inference [9],[10], which the present study further confirms in the specific context of G0 seed potato sorting. Likewise, the substantial gain in mAP@0.5–0.95 observed between epoch 25 and epoch 100 echoes the architectural improvements documented for recent YOLOv11-based agricultural detectors, such as DMN-YOLO for apple leaf disease detection and YOLO-MECD for citrus detection [19],[20], both of which attribute performance gains to enhanced backbone and attention mechanisms rather than dataset size alone, supporting the interpretation that the compact C3k2 and C2PSA components of YOLOv11n are primarily responsible for the strong generalization observed in this study despite the limited training data.

From the perspective of AIoT integration, the present findings align with and advance the direction set by previous deep-learning-and-IoT convergence studies. The real-time crowd counting system that combined YOLOv8 NAS with IoT sensors achieved an mAP of 95.1% with sub-two-second response latency [17], while the IoT-based fish seed counting prototype reported a maximum counting error of only 3.65% [18]; both studies demonstrate that pairing deep-learning-based detection with lightweight IoT communication can deliver accurate, near-real-time monitoring, a conclusion that the present study reinforces through the 0.986 mAP@0.5 achieved alongside fully synchronized MQTT-based dashboard visualization. However, unlike these earlier systems, which were restricted to single-category counting without size-based discrimination, the proposed system extends this line of research by simultaneously performing multi-category size classification, object counting, and real-time remote monitoring within a single integrated pipeline, an integration gap that was also identified, but not directly addressed, in the comprehensive AIoT framework proposed for crop yield prediction [8]. This comparative positioning is consistent with the benchmarking presented in Table 2, where the modified SCG-YOLOv8n potato counting framework [12], the YOLOv8-ByteTrack maize kernel counting system [14], and the YOLOv8-HD wheat seed detector [15] each achieved strong detection or counting metrics in isolation but did not incorporate IoT-based remote synchronization or size-grading functionality, reaffirming that the principal contribution of this study lies not merely in detection accuracy but in the unified convergence of accurate visual classification with reliable IoT-based data transmission for G0 seed potato production monitoring.

4. Conclusions and Future Works

This study successfully developed and validated a YOLOv11n-based AIoT system for automated size detection and counting of G0 seed potatoes using the MQTT protocol. The system integrated a conveyor belt, USB webcam, and laptop as the edge computing unit, with detection results transmitted in real-time to a Node-RED dashboard via MQTT broker. The YOLOv11n model was evaluated across four epoch configurations, with epoch 100 identified as the optimal setting, achieving precision approaching 1.00, recall of approximately 0.98, mAP@0.5 of 0.986, and mAP@0.5–0.95 of 0.96, the highest values among all tested configurations. The progressive improvement in mAP@0.5–0.95 from 0.91 at epoch 25 to 0.96 at epoch 100 confirms that extended training meaningfully enhances bounding box localization accuracy, which is directly critical for the pixel-area-based size classification task. End-to-end validation across 30 trials demonstrated a perfect accuracy of 100% in both total potato counting and per-category size classification, confirming the reliability of the complete AIoT pipeline and its suitability for automated G0 potato production monitoring.

For future works, it is recommended to compare YOLOv11n against other object detection architectures, such as SSD and Faster R-CNN, under identical computational conditions. Additionally, exploring instance segmentation-based size measurement methods, such as Mask R-CNN, could improve boundary-mask classification accuracy. Most importantly, to address critical environmental boundary conditions, future research must prioritize long-term real-field validation within operational greenhouse environments. This includes stress-testing the system under highly variable natural illumination, dynamic outdoor weather dependencies, high dust accumulation on camera lenses, diverse crop morphologies, and significantly higher conveyor throughputs. Such field deployments are essential to comprehensively assess the generalizability, environmental resilience, and industrial scalability of the proposed AIoT system for commercial-scale G0 potato seed production in Indonesia

5. References

[1] FAO, *World Food and Agriculture - Statistical Yearbook 2022*. Rome, 2022. doi: 10.4060/cc2211en.

[2] Kurniawati, A. Nikmatullah, and Kisman, "Respon Pertumbuhan dan Hasil Empat Varietas Kentang Industri (Solanum tuberosum L.) pada Dua Musim Tanam Berbeda di Desa Sajang, Sembalun, Lombok Timur," *J. Ilm. Mhs. Agrokomplek*, vol. 4, no. 1, pp. 112–117, 2025, doi: 10.29303/jima.v4i1.6229.

[3] R. Wardana, F. L. Afriansyah, and H. Oktafa, "Revitalization of the Potato Center through the Implementation of Smart Farming and Smart Agrology Systems in the Production Process and Development of Mashed Potato products in Bondowoso," in *7th National Conference for Community Service (NaCosVi)*, Politeknik Negeri Jember, 2024, pp. 378–382.

[4] A. Sembiring, R. Murtiningsih, J. P. Sahat, and S. Hartanto, "GO potato aeroponic seed production in Indonesia, producers' perception toward the benefits and challenges," in *IOP Conference Series: Earth and Environmental Science*, IOP Publishing Ltd, 2021. doi: 10.1088/1755-1315/948/1/012086.

[5] S. Deras, "peningkatan efisiensi agribisnis kentang melalui introduksi bibit Granola 4 di Desa Bandar Hinalang Kecamatan Purba Kabupaten Simalungun," *J. Agriust*, vol. 1, pp. 15–22, 2020, doi: 10.54367/agriust.v1i1.1023.

[6] B. L. Candradewani, S. Indriyanto, and I. Permatasari, "Sistem Monitoring Kelembapan Media Tanam Aglaonema SP Berbasis Internet of Things (IoT)," *J. SINTA Sist. Inf. dan Teknol. Komputasi*, vol. 2, no. 3, 2025, doi: 10.61124/sinta.v2i3.60.

[7] M. Geovanie, I. Ruslianto, and U. Ristian, "Sistem Pemantauan dan Kendali Tanaman Kentang Media Aeroponik Berbasis Internet of Things," *CESS (Journal Comput. Eng. Syst. Sci.*, vol. 8, no. 1, 2023, [Online]. Available: http://www.jurnal.unimed.ac.id

[8] M. S. Sharafat and others, "An IoT-enabled AI system for real-time crop prediction using soil and weather data in precision agriculture," *Smart Agric. Technol.*, vol. 12, p. 101263, 2025, doi: 10.1016/j.atech.2025.101263.

[9] W. Sun, X. Niu, Z. Wu, and Z. Guo, "Lightweight Detection Counting Method for Pill Boxes Based on Improved YOLOv8n," *Electronics*, vol. 13, no. 24, p. 4953, 2024, doi: 10.3390/electronics13244953.

[10] J. Peng and others, "Towards toxic and narcotic medication detection with rotated object detectors," *Quant. Imaging Med. Surg.*, vol. 13, no. 4, pp. 2156–2166, 2023, doi: 10.21037/qims-21-1146.

[11] D. H. Saputra, B. Imran, and J. Juhartini, "object detection untuk mendeteksi citra buah buahan menggunakan metode yolo," *J. Kecerdasan Buatan dan Teknol. Inf.*, vol. 2, pp. 70–80, 2023, doi: 10.69916/jkbti.v2i2.18.

[12] L. Du, J. Li, W. Yang, W. Liu, X. Sun, and Y. Xia, "A SCG-YOLOv8n potato counting framework with efficient mobile deployment," *Sci. Rep.*, vol. 15, no. 1, p. 34909, 2025, doi: 10.1038/s41598-025-18754-9.

[13] A. M. Iqbal, J. Jumadi, and E. Nurlatifah, "Implementasi YOLOv8 Sebagai Pendeteksi Nominal Uang Rupiah Kertas Berbasis Android," *SMATIKA J.*, vol. 15, no. 02, pp. 268–277, 2025, doi: 10.32664/smatika.v15i02.1545.

[14] R. Li and others, "Maize Kernel Batch Counting System Based on YOLOv8-ByteTrack," *Sensors*, vol. 25, no. 17, p. 5584, 2025, doi: 10.3390/s25175584.

[15] N. Ma, Y. Su, L. Lee, Z. Li, and H. Yan, "Wheat Seed Detection and Counting Method Based on Improved YOLOv8 Model," *Sensors*, vol. 24, no. 5, p. 1654, 2024, doi: 10.3390/s24051654.

[16] S. M. Alhammad, D. S. Khafaga, W. M. El-hady, F. M. Samy, and K. M. Hosny, "Deep learning and explainable AI for classification of potato leaf diseases," *Front. Artif. Intell.*, vol. 7, pp. 1–9, 2025, doi: 10.3389/frai.2024.1449329.

[17] S. K. Jayasingh, P. Naik, S. Swain, K. J. Patra, and M. R. Kabat, "Integrated Crowd Counting System Utilizing IoT Sensors, OpenCV and YOLO Models for Accurate People Density Estimation in Real-Time Environments," in *2024 1st International Conference on Cognitive, Green and Ubiquitous Computing (IC-CGU)*, IEEE, 2024, pp. 1–6. doi: 10.1109/IC-CGU58078.2024.10530804.

[18] S. Gunawan, A. E. Setiawan, and D. Feriyanto, "Design of Prototype Gourami Fry Counting System Based on Internet of Things (IoT) Using Infrared Sensor," *Int. J. Softw. Eng. Informatics*, vol. 1, no. 1, pp. 40–48, 2023, [Online]. Available: https://journal.aisyahuniversity.ac.id/index.php/IJosei

[19] L. Gao, H. Cao, H. Zou, and H. Wu, "DMN-YOLO: A Robust YOLOv11 Model for Detecting Apple Leaf Diseases in Complex Field Conditions," *Agriculture*, vol. 15, no. 11, p. 1138, May 2025, doi: 10.3390/agriculture15111138.

[20] Y. Liao, L. Li, H. Xiao, F. Xu, B. Shan, and H. Yin, "YOLO-MECD: Citrus Detection Algorithm Based on YOLOv11," *Agronomy*, vol. 15, no. 3, p. 687, 2025, doi: 10.3390/agronomy15030687.

[21] I. Permatasari, N. A. F. Tanjung, and N. A. Zen, "Perancangan Sistem Monitoring Konduktivitas dan Padatan Terlarut PDAM Banyumas Berbasis IoT," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 10, no. 1, pp. 25–31, 2021, doi: 10.22146/jnteti.v10i1.1023.

[22] Y. Li, "An Enhanced YOLOv11 for Small-Object Detection in UAV Imagery," in *Proceedings of the 2025 8th International Conference on Computer Information Science and Artificial Intelligence*, ACM, 2025, pp. 574–579. doi: 10.1145/3773365.3773458.

[23] R. Van Glabbeek, D. Deac, T. Perale, K. Steenhaut, and A. Braeken, "Flexible and Efficient Security Framework for Many-to-Many Communication in a Publish/Subscribe Architecture," *Sensors*, vol. 22, no. 19, p. 7391, 2022, doi: 10.3390/s22197391.