Journal of Innovative Image Processing (ISSN: 2582-4252)
www.irojournals.com/iroiip/

Instance Segmentation for Local Rice Seed Germination Evaluation using Deep Learning

Diane B. Remot[1], Cherry R. Gumiran[2]

[1,2]Information Technology Department, College of Computing Studies, Isabela State University-Cauayan Campus, Cauayan City, Isabela, Philippines.
[1]Information Technology Department, Eastern Visayas State University-Main Campus, Tacloban City, Leyte, Philippines.

E-mail: [1]diane.remot_cyn@isu.edu.ph, [1]diane.remot@evsu.edu.ph, [2]cherry.r.gumiran@isu.edu.ph

Abstract

This study combines sophisticated deep learning technology to solve the manual procedure of rice seed germination evaluation which is time-consuming and error-prone. This paper examines the architectural behavior of CNN architectures in detecting, segmenting and classifying local rice seeds into normal, abnormal, fresh ungerminated and dead seeds. This is done using Convolutional Neural Network (CNN) architectures such as You Only Look Once (YOLO)v8s-Seg and YOLOv9c-Seg. A high-end smartphone was used to capture the image in natural light, and polygon annotation was used to label an image of seeds that covers individual, partially occluded and overlapping seeds in a single image. A post-processing technique called Segment Anything Model (SAM) was adopted to improve the mask borders of the rice seed objects for accuracy. The outcome illustrates that both models function well with a mean Average Precision (mAP@50) of more than 90%. However, there are misclassification cases identifying rice seeds between dead and fresh ungerminated, and then abnormal and normal. The YOLOv8s-Seg was shown to be more computationally efficient for mobile deployment where it has 40.2 Giga Floating Point Operations Per Second (GFLOPs). Meanwhile, YOLOv9c-Seg was found to be better for feature extraction and fast convergence. However, YOLOv9c-Seg was computationally complex for mobile edge deployment. The research reveals that the incorporation of automated seed quality assessments from an artificial intelligence (AI) framework to empower the process can drastically ramp up agricultural productivity and sustain economic growth.

Keywords: YOLOv8s-Seg, YOLOv9c-Seg, Rice Seed Germination, Segment Anything Model (SAM), Instance Segmentation, Rice Seed Quality Assessment, Detection.

1. Introduction

Integrating various technologies through innovation and research is already part of the growing nature of agriculture. From [1], it shows how agricultural methods are changing as a result of smart farming technologies, including IoT, sensors, artificial intelligence, and driverless cars. Critical global issues like population increase, food security, and climate change are the focus of these technologies. However, significant barriers remain among farmers as

DOI: https://doi.org/10.36548/jiip.2026.1.011
Received: 16.12.2025, received in revised form: 13.01.2026, accepted: 27.01.2026, published: 07.02.2026

implementation costs, internet connectivity issues, and a lack of technical knowledge are common challenges to realizing development and its application [2]. With this, a continued exploratory study is vital to secure progressive and precision farming.

Food consumption is essential for human survival, and rice is a staple in every main dish in the Philippines. Not only does it satisfy hunger, but it also replenishes energy for the day ahead. The quality of rice is crucial for providing consumers with delicious and nutritious grains. In the Philippines, the Bureau of Plant Industry (BPI) operates under the Department of Agriculture (DA) and is primarily responsible for promoting and regulating the country's plant industry, including crop production.

To maintain competitiveness and sustainability as a rice producer, various seed growers in the Philippines may submit their local rice samples to the Bureau of Plant Industry through its Seed Certifying Division, the National Seed Quality Control Services Division (NSQCSD), for quality assessment and tagging as certified seed. There are four mandated evaluation processes to assess rice seed quality: moisture content analysis (for storage), physical tests (to ensure the absence of seed-borne diseases), varietal tests (for genetic characteristics), and the germination test (physiological). By using high-quality seeds, healthier crops that grow faster can positively contribute to environmental changes, thereby meeting market demands and consumer needs.

To foster a fast and accurate approach to rice seed quality evaluation during the germination phase, the application of appropriate technologies like artificial intelligence can support the current manual task, which is time-consuming in counting germinated and non-germinated rice seeds and is prone to human errors. Several studies use Convolutional Network Models (CNN)[3][4] for various rice seed assessment tasks and demonstrate their potential to transform rice seed assessment by providing a fast, accurate, and automated evaluation method that can significantly improve agricultural productivity. To reduce workload and support the challenging rice assessment task in the germination phase, the researchers aim to explore the morphological characteristics of rice seeds after the germination period using CNN technology, such as advanced deep learning models like YOLOv8s-seg and YOLOv9s-seg. The performance metric of both models may be able to evaluate their accuracy in identifying the rice seed status after germination into four classes: normal, abnormal, fresh ungerminated, and dead seed, and determine which models are suitable for edge deployment on mobile applications. There are several studies that delved into this, but not with multiclass identification or detection that includes partially occluded or overlapping seeds of local rice, which addresses the gap that this study will explore.

2. Related Work

Several approaches to various research in agriculture are relevant in contributing to sustainable agricultural progress, which includes the integration of various technologies that are adaptable and practical.

Existing machine learning studies use deep learning models to improve processes in agriculture and are implemented to achieve economic outcomes that will likely support the sustainable growth of agricultural products such as rice. The use of advanced deep learning models, such as YOLO (You Only Look Once), can provide a relevant interpretation of what humans can detect. Authors [5] report that the RSG-Yolo model can accurately detect and classify rice seeds as germinated or non-germinated, while [6] shares a method for detecting the seed germination rate of wild rice using SGR-Yolo, achieving an accuracy rate of 94% when tested in a hydroponic box and 98% in a petri dish. The difference between wild rice and local rice lies in their context of preparedness and location of growth. Wild rice is a more specialized choice due to its nutritional density, promoting healthier consumption, while local rice is less expensive and easier to prepare. Yolov8 was an effective model used to identify and separate the endosperm of seeds, which is crucial for efficient crop planting, through X-ray morphological analysis [7]. A study [8], which includes feature segmentation and extraction during germination, improves detection and segmentation by 1 to 2% using YOLOv8-segANDCAL on soybean radicle features. Another study [9] classifies germinated seeds as normal, abnormal, or ungerminated for pole sitao using SSD MobileNet and Faster R-CNN models, which could potentially expedite and improve reliability in predictions. To detect germinated corn seeds [10], a YOLOv8n model was used to enhance target detection and complexity, making it suitable for efficient predictions.

This YOLOv8 model and its segmentation feature were also utilized in these studies to detect licorice seeds [11], segment healthy and diseased tomato plants [12], and analyze and quantify pea seeds [13] in their germination state. It has been found effective in measuring vigor to produce more crops and plants, leading to a healthier agricultural industry. The study [14] aimed to compare two models, YOLOv8 and Mask R-CNN, under instance segmentation of orchard environments, where YOLOv8 outperformed the other models in terms of suitability for developing smart and automated orchard operations for real-time applications, such as robotic farming. The YOLOv5 model was used to detect real-time plant health, checking if plants are healthy or diseased so that farmers can take necessary precautions to halt disease spread [15]. From the study [16], the detection of weeds and classification of their growth stages demonstrate that YOLOv5 and RetinaNet perform better compared to Faster R-CNN, with the highest recall at 79.4% for YOLOv5 and a precision of 87.46% for RetinaNet.

The presence of these various studies using YOLO (You Only Look Once) models explicitly describes their relevance to applications in agricultural processes. The contribution of this work implies that using models YOLOv8s-Seg and YOLOv9c-Seg may yield suggestive and favorable results that can detect, segment, and classify local rice seeds as normal, abnormal, fresh ungerminated, and dead based on their morphological characteristics, distinguishing them after a seven-day germination period. The results of this study may support the time-consuming task of BPI-NSQCSD or an agency needing this, mitigating error-prone identification and counting of local rice seeds. Studies have indicated that behaviorally focusing on the detection and segmentation of multiclass characteristics of germinated rice seeds in individual, partially occluded, and overlapping scenarios has not yet been fully realized in the existing literature. It is timely to conduct this research using the models to investigate how they behave in appropriately identifying germinated rice seeds as normal, abnormal, fresh ungerminated, and dead.

3. Proposed Work

This research will use integrated technology employing computer vision to detect, segment, and classify local rice seeds into four classes: normal, abnormal, fresh ungerminated, and dead, including partially occluded and overlapping seeds. The individual and mixed (individual, occluded, or partial overlap) images were preprocessed and then labeled using computer vision software such as the Roboflow online platform and annotated with a polygon tool. Data preparation is crucial, as the interpretation of the model depends on how it was preprocessed and labeled. The images of rice samples come from various variants of local rice seeds available in NSQCSD and approved by the Philippine Seed Board (PSB) and the National Seed Industry Council (NSIC), coded as RC 216, RC 222, RC 402, RC 422, RC 442, RC 506, and RC 508.

There are 400 rice seeds per variant submitted and processed for germination in 7 days, and separated per replication of a 100 rice seeds on filter paper, as shown in Figure 1. The 4 replications were sprayed with distilled water, then rolled over and placed in a clear plastic bag to maintain their moisture before transferring them to the germinator at an appropriate temperature of 30 degrees in the morning and 20 degrees in the evening.

Figure 1. Rice Seed Spread on a Filter Paper in One Replication

After a 7-day germination period, the rice seeds were removed from the germinator, and the numbers of normal, abnormal, fresh ungerminated, and dead seeds were assessed to determine if they met the 85% passing germination rate. To meet this passing rate, each variant must achieve 85% or more of the normal rice seeds across the four replications. If any variant fails to meet the required rate, an extended seven (7) day germination period will be repeated, but the abnormal and fresh ungerminated seeds will go through the process again, as they still have the potential to germinate. Figure 2 shows the results of the germinated rice seeds.

Figure 2. Germinated Rice Seed after a Seven (7) Day Germination Period

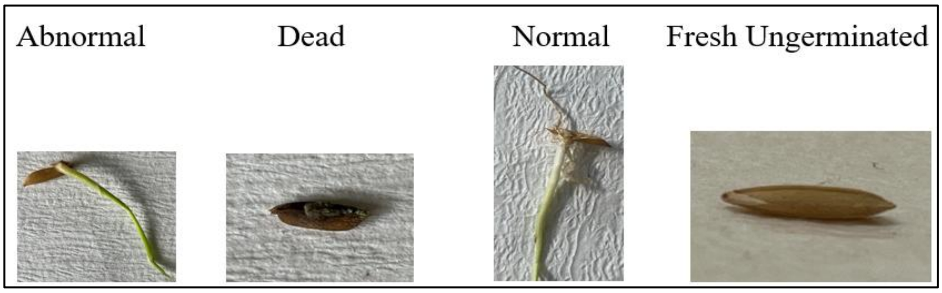

Figure 3. Classes of Rice Seed (Raw Image)

For the models to detect, segment, and classify a rice seed as normal, abnormal, fresh ungerminated, or dead, as shown in Figure 3, its morphological characteristics play a vital role. Abnormal seeds exhibit incomplete, deformed, or damaged germination organs that hinder further development but have the potential for continued growth during an extended germination period of another seven days after the first count. A dead rice seed shows no signs of seedling development and exhibits discolored and moldy characteristics; hence, this type cannot develop or germinate under any conditions. A fresh ungerminated seed is firm, has a normal color, shows no signs of decay, and lacks developing roots and shoots, but can be re-germinated for an extended period of seven days after the first count. A normal seed has a well-developed and intact root and shoot structure capable of sustained growth.

Figure 4. Techniques Applied for Rice Seed Preprocessing (Individual Image)

The dataset preparation consumes time and lengthens the task since image capture is done through an iPhone 13 smartphone under natural light to get the images' raw texture of 3,328, avoiding shadows. It is first-hand data that is already publicly available in Roboflow, and labeling protocols are provided to ensure reproducibility. There were 16,630 images preprocessed, and a sample shown in Figure 4 applying techniques such as Gaussian blur and skeletonization using the Zhang–Suen thinning algorithm preserves the overall shape and topological structure of rice seeds, thereby improving the instance segmentation performance, especially under occlusion conditions. Removing the background and segmenting images enhances the clean images, avoiding later noise that will distract the model from detecting. These steps ensured that the images were cleaned and able to suppress high-frequency noise before labeling and training, as seen in Figure 5, where the outcome looks the same as the raw images.

Figure 6 represents a preprocessing technique applied to raw images in one frame from various classes. The refinement of images contributes to reducing image noise while preserving the edges of the rice seed to which the Gaussian blur was applied. In improving the local contrast of rice seeds, the lab color space L-Channel(lightness) isolation and application of Contrast Limited Adaptive Histogram Equalization (CLAHE) preserve color relationships, which is crucial for rice seeds, as it helps the model distinguish subtle surface defects or any abnormal features that might be washed out in standard RGB lighting. The merged result will give a sharper distinction between the seeds, especially fresh ungerminated, and dead seeds, and their background, making it more precise during training.

Four classes of rice seeds were then labeled or annotated using Roboflow, an online computer vision software tool that is necessary for detection and segmentation, allowing models to read points in the image. A polygon tool was utilized to ensure the shape and segmentation of the rice seed per class, as it is elongated, its shoots and roots are curved, and shapes are non-rectangular and irregular, as seen in a sample in Figure 7. However, 154 images were not included in the annotation due to image capture, and some seeds were deemed blurry and incomplete in terms of seed shape.

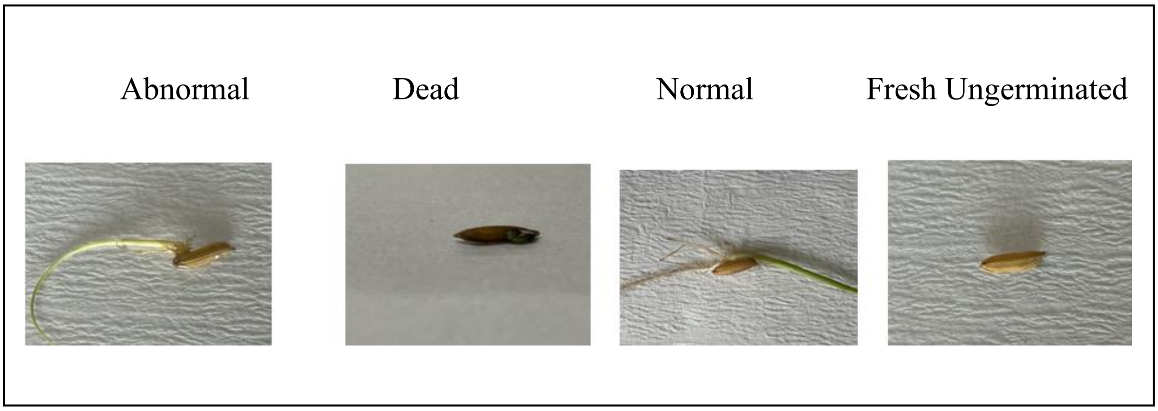

Figure 5. Sample of a Preprocessed Rice Seed (Individual Image)

Figure 6. Sample of a Preprocessed Rice Seed (Multiple Images)

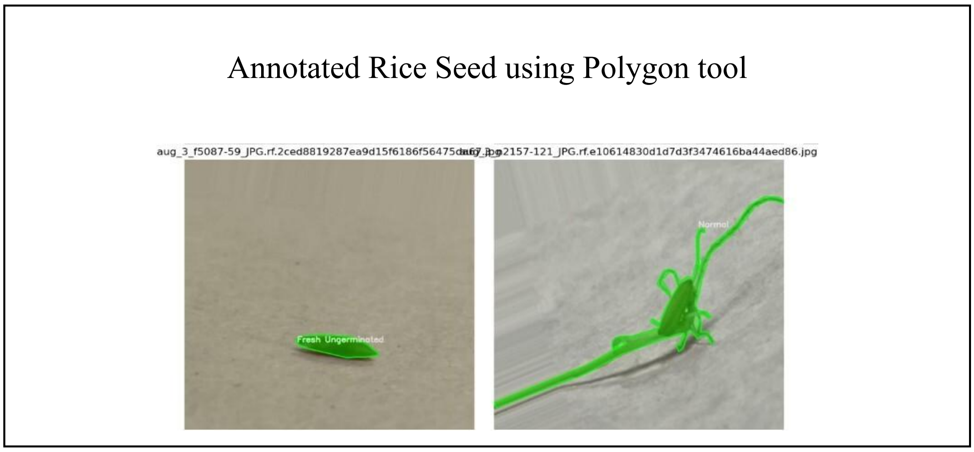

Figure 7. Labeled Sample of Rice Seed in Individual and Multiple Images

Figure 8 illustrates the workflow of this study from data preparation, data transformation, training, and data visualization. A comprehensive computer vision pipeline for real-time instance segmentation is demonstrated by this method. It starts with image capture from a mobile device. Roboflow's manual annotation and data preparation create high-quality ground truth labels that are necessary for accurate training.

Through careful splitting at the image level of a dataset, with a 70% training, 20% validation, and 10% test split of annotated images made before augmentation requirements were applied to the training set, it ensures that all seed instances within an image appear in one split, allowing for data normalization. The transformation phase gets the data ready for a reliable model training. To facilitate comparative performance analysis, the core data mining utilizes two advanced YOLO architectures: YOLOv8s-seg for efficiency and YOLOv9c-seg for improved speed and feature extraction. However, there could be cases where variations in the morphological characteristics of rice seeds after the germination period, could affect detection and segmentation accuracy, similar to the challenge noted in [17]. The data preparation can initially affect the outcome, but can be managed according to a series of test simulations. Through a variety of output formats, including segmented object masks, detection accuracy metrics, and model comparison dashboards, the visualization phase offers essential details. Accordingly [12], algorithm performance has increased, as a result of YOLOv8's substantial enhancements and optimizations over the YOLOv5 network. In addition to object detection and tracking, the YOLOv8 network can perform other tasks, such as instance segmentation, image categorization, and key point detection. The YOLOv8 object detection paradigm is expanded upon by YOLOv8s-Seg, which is specifically designed for segmentation tasks. The YOLOv8s-Seg network maintains a high segment mean average precision while achieving real-time object instance segmentation by utilizing the concepts of the YOLACT network, which is timely to test in the analysis of multi-class germinated rice seeds. The YOLOv8s-seg model in this study was trained at a resolution of 768x768 for 80 epochs using boundary-focused augmentations and adjusted loss weights for individual rice seeds. This enables learning in a boundary-focused area of the rice seed for normal and abnormal rice seeds with germination structures. For images of mixed individual and occluded rice seeds, YOLOv8s-Seg employed balanced loss weighting and overlap-aware mask settings with controlled mosaic augmentation.

Segmentation models in deep learning are being used increasingly in agricultural applications. For example, YOLOv9-based detection models have shown strong seedling and weed detection for soybean crops [18]. The YOLOv9-C variant has demonstrated improved feature extraction capabilities in seed quality tasks and has been effectively applied to surface defect identification in soybean seeds [19]. To take advantage of its improved gradient propagation and feature retention for challenging morphology segmentation tasks, YOLOv9c-seg, which contains a Generalized Efficient Layer Aggregation Network (GELAN) lightweight network architecture, may be suitable for further testing on multi-class germination rice seeds. It uses higher-resolution inputs with specified color and geometric augmentations and segmentation-specific settings to properly capture the shape of a rice seed in individual images. It applies a stronger early-stage mosaic augmentation and better geometric and segmentation settings so that the model can analyze objects in occluded conditions.

These approaches enable deployment optimization and iterative model improvement, which are beneficial for mobile-based applications that require real-time segmentation capabilities. Through model comparison and performance validation, the pipeline is open to enhancing the model's capabilities, with a focus on practical deployment concerns. The model's performance assessment in this study was tested in an offline inference using pre-captured images.

The dependency of results requires the training parameters, such as hardware, data acquisition device, and software, to visualize the patterns that the models were interpreting and learning, as shown in Tables 1 and 2, respectively.

Figure 9 provides a number of image data used for training simulation, including the validation and test sets, where a balanced class weight was provided to overcome overfitting.

Each image had instances per class, which is normal for a YOLO model, and once an image was labeled, it yielded abnormal and normal seeds as high as 1.12 and 1.14 seeds per image, respectively, due to the boundary complexity in their characteristics, based on their roots and shoots.

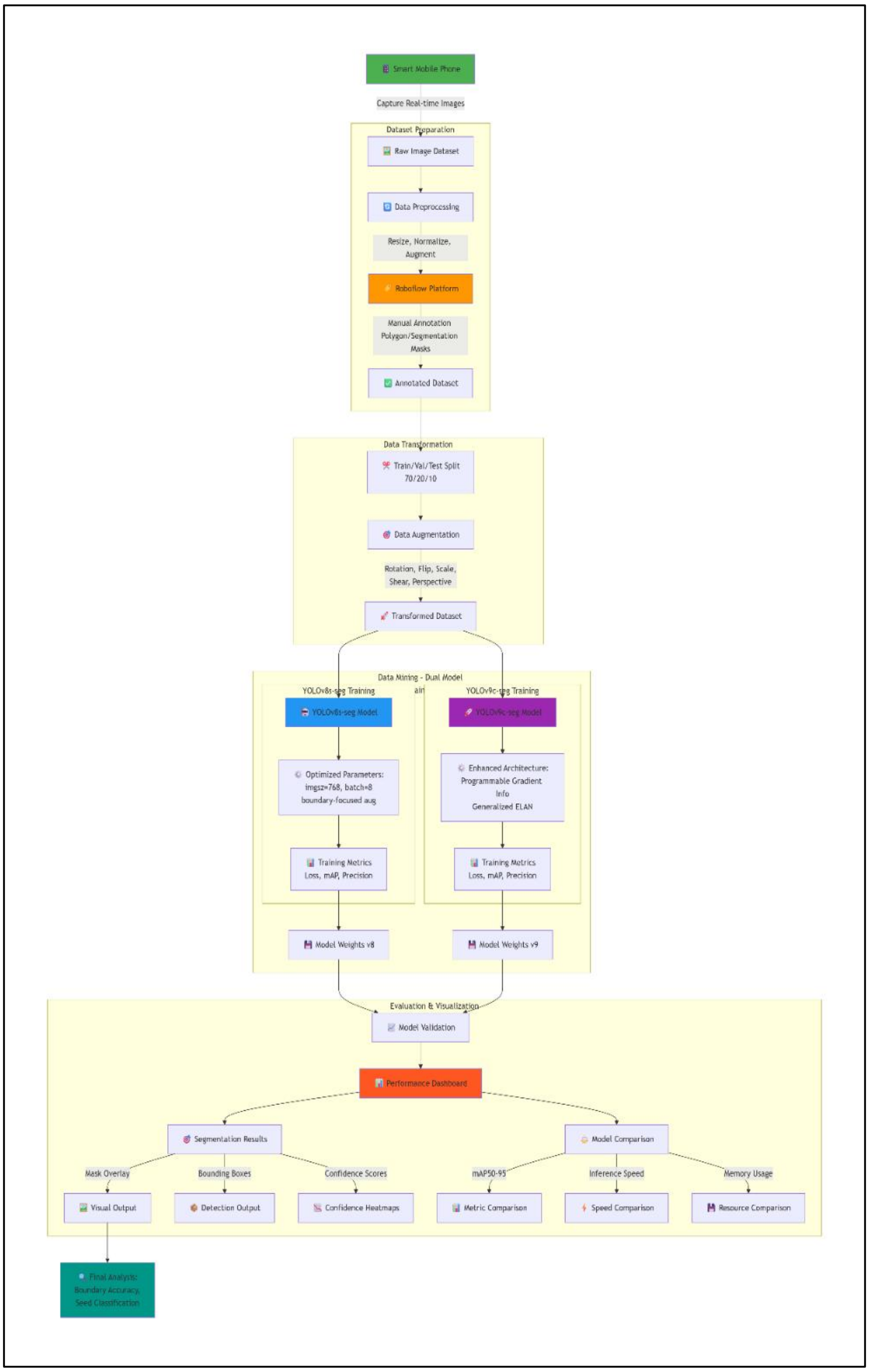

Figure 8. A Comprehensive Machine Learning Pipeline for Rice Seed Germination Analysis Using Models YOLOv8s-Seg and YOLOv9c-Seg

Compared to the other two classes, fresh ungerminated and dead seeds had their shape and color as a basis for measuring instances, with 1 seed per image equivalent to 1 instance.

Figure 10 provides dataset-annotated images to be used in training with mixed individual and multiple images, so the model can learn effectively from its ground truth based on its validation and test sets. A balanced class weight formula was obtained to address class imbalance, where the normal rice seed is in the majority.

3.1 Training Requirements

Table 1. Hardware and Data Acquisition Device (DAD)

Hardware			**Data Acquisition**
CPU	GPU	RAM	Mobile Device
AMD Ryzen 7 5700G with Radeon Graphics (3.80 GHz)	NVIDIA GeForce RTX 3060 12 GB	32.0 GB (31.8 GB usable)	Smart Mobile Phone (iPhone 13); Lighting: Natural Light

Table 2. Software

Training Platform/Output	**Annotation**
Jupyter Notebook	Roboflow

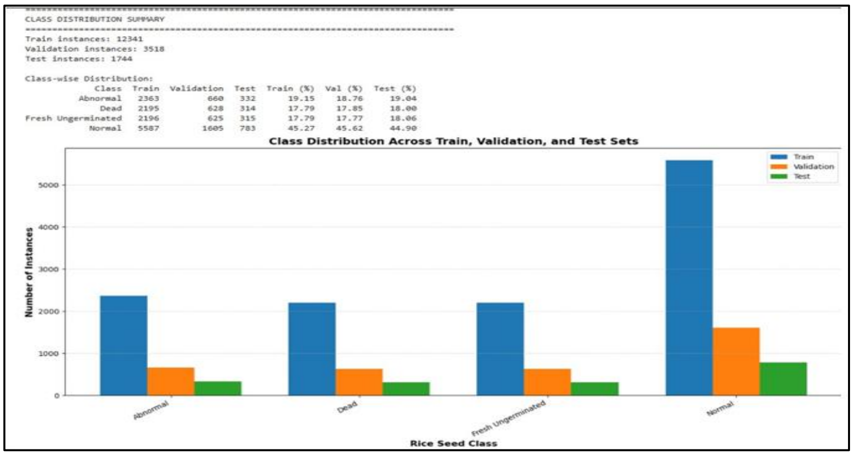

Figure 9. Dataset Class Distribution (for Individual Image)

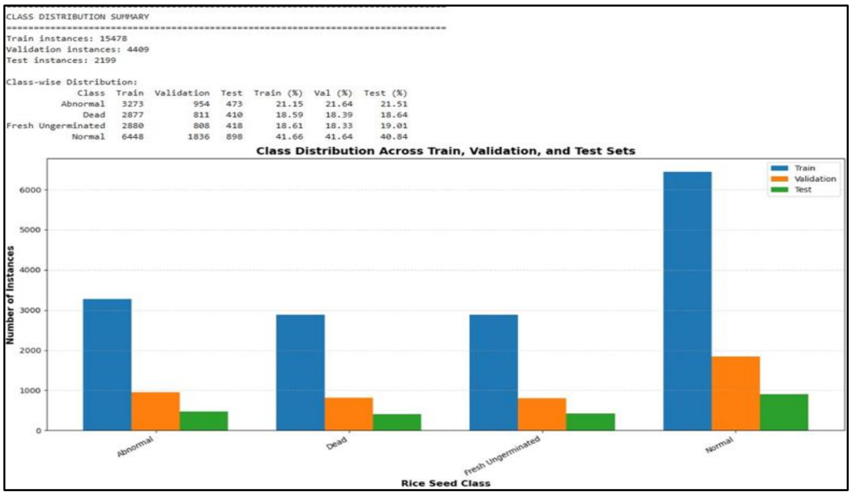

Figure 10. Dataset Class Distribution (for Mixed Images)

Figure 11 shows the architecture of a YOLOv8s-seg model tested on individual and mixed image datasets. For the simultaneous detection, classification, and instance segmentation of germinated rice seeds, the YOLOv8s-seg model provides an efficient end-to-end framework. It achieves a good balance between precision and processing efficiency, for a standard image size of 640 pixels, it produces 261 layers comprising 11.8M parameters and 40.2 Giga Floating-Point Operations (GFLOPs), making it suitable for both agricultural imaging and edge deployment, such as mobile. The three major blocks of this model include a backbone, neck network, and segmentation head, which contribute to efficient performance. Mentioned in the study of [20], its backbone network is composed of a series of convolutional layers, C2f blocks, and a Spatial Pyramid Pooling–Fast (SPPF) module for hierarchical and multi-scale feature extraction of rice seed images that can simulate subtle morphological variations associated with germination stages, including radicle development, deformation, and surface roughness, without significantly increasing model complexity due to the C2f blocks' enhanced feature reuse and gradient propagation. Part of the model's strong performance is its PAN-FPN neck, which aggregates multi-scale features of rice seeds and improves detection and segmentation performance across objects of varying sizes. This is particularly applicable in scenarios where objects overlap, as it captures fine spatial details that contribute to the overall form and pattern. The segmentation head provides a prototype-based mask generation mechanism that jointly performs bounding box regression, class prediction, and instance mask generation, enabling accurate detection and segmentation of rice seeds into normal, abnormal, fresh ungerminated, and dead classes. This makes the model an appropriate approach for agricultural and image industry analysis, due to its balanced accuracy and computational efficiency.

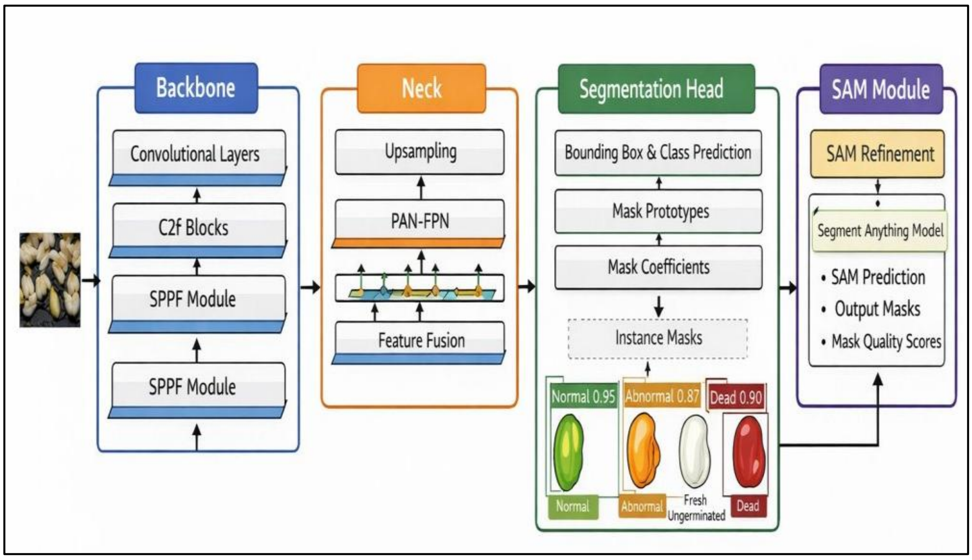

Figure 11. YOLOv8s-Seg with Segment Anything Model (SAM)Rice Seed Germination Architecture

The integration of the SAM model in Figure 11 provides additional information on the accuracy of detection of YOLOv8s-seg, which enables efficient results specifically for raw or real-time image capture. This can elevate the stricter IoU mask provided by the YOLO model after training to refine segmented rice seed boundaries and harmonize with YOLOv8s-seg detection on partially touching, occluded, or overlapping seeds, allowing for local separation and classification accordingly.

Figure 12 simplifies the YOLOv9c-Seg model process in treating raw mixed images from various rice seed classes, dealing with pre-processing, training, and post-processing techniques to detect, classify, and segment accordingly. It has 654 layers, 27.8M parameters, and 149 GFLOPs, which makes this model computationally complex for mobile deployment but yields better results in detection and segmentation, particularly when the standard image size is increased from 640 pixels to 897 pixels. The process begins with the backbone network to extract spatial and semantic features of an input rice seed image to uncover the morphological characteristics for the model to learn. In the head network, these extracted features enable the model to detect, classify, and segment individual seeds simultaneously through a multi-scale feature fusion technique, where shallow (edges, textures, and small shapes) and deep layers (object type and overall shape) of objects are merged together. The presence of a post-process technique through the Segment Anything Model (SAM) refines the coarse mask generated from the initial segmentation, improving boundary accuracy and the separation of partially touching, occluded, or overlapping seeds.

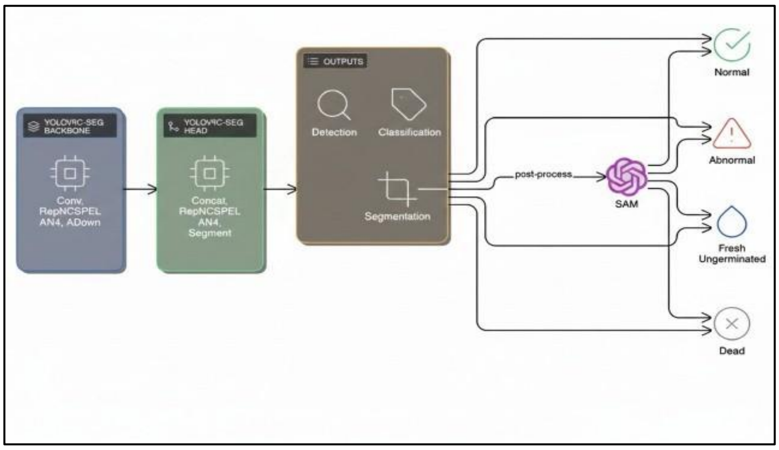

Figure 12. YOLOv9c-Seg with Segment Anything Model (SAM)Rice Seed Germination Architecture

4. Results and Discussion

This section evaluates the validity of each model's behavior in detecting, segmenting, and classifying multiple classes of local rice seeds. The results will provide input for an appropriate application for real-time image capture using a mobile device, ensuring the model's detection capability is suitable for mobile deployment.

4.1 YOLOV8s-Seg and YOLOV9c-Seg Model Performance Metrics (for Individual Dataset)

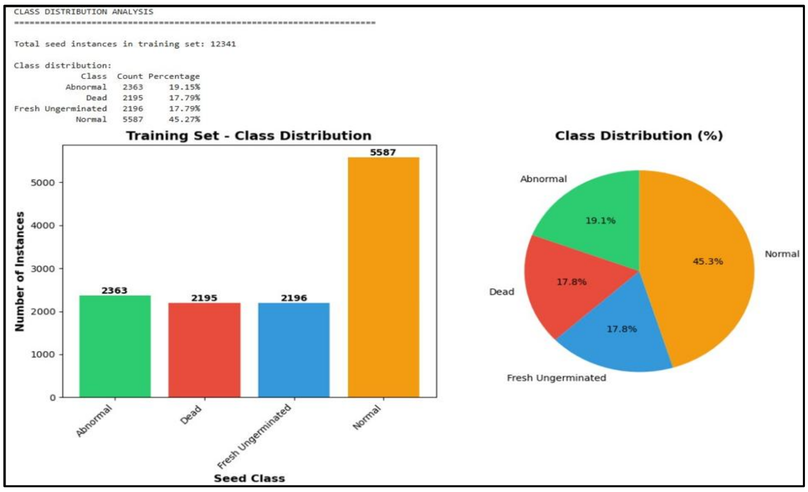

Figure 13. Training Dataset Class Distribution (Individual Image)

In YOLO models, an image that is labeled has an equivalent instance. There could be a case where one seed is detected with 2-3 objects or instances. A higher instance count is a good sign of better object diversity and stronger learning of class boundaries. Figure 13 provides 12,341 instances from 11,501 images, resulting in 0.93 instances per image, indicating that most of the seeds contain one object, with a few having multiple seeds. It also displays the training dataset representing the multi-class rice seed as a reference for the simulation. The distribution of class weights balances the rest of the classes in normal, abnormal, fresh ungerminated, and dead seeds before training to avoid overfitting or underfitting.

Figure 14 shows the recall performance of the two models on both the validation and test sets for an individual image. A consistent result of high recall from both models was observed, enabling effective detection and minimizing the presence of false negatives. YOLOv9c-seg slightly outperforms YOLOv8s-seg based on the validation set, with a recorded 92.7% recall for box detection and 93.1% for the segmentation mask. In the test set, YOLOv8s-seg indicates better box and segmentation recall, slightly outperforming YOLOv9s-seg. The outcome clearly indicates that both models, based on ground truth, can successfully detect, classify, and segment the actual rice seeds in the germination analysis with minimal false negatives.

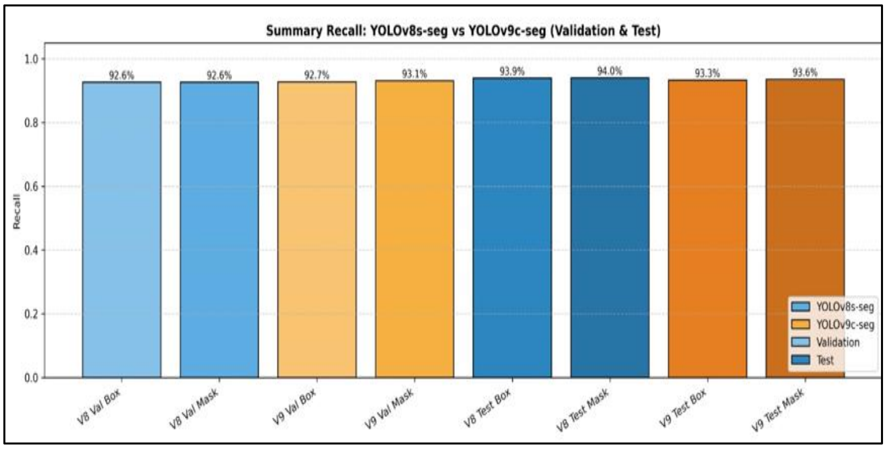

Figure 14. Recall Summary (Comparison of YOLOv8s-Seg and YOLOv9c-Seg Performance Metrics on Its Validation and Test Set)

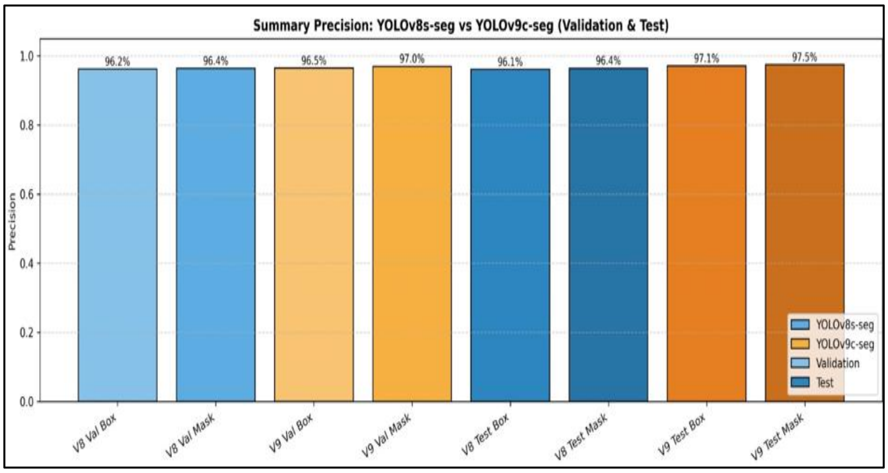

Figure 15. Precision Summary (Comparison of YOLOv8s-Seg and YOLOv9c-Seg Performance Metrics on Its Validation and Test Set)

Figure 15 shows the models learning to predict the actual class of an image as normal, abnormal, fresh ungerminated, and dead seed. Both in the validation and test set, YOLOv9c-Seg slightly predicts better than YOLOv8s-seg.

The standard metric for YOLOv8 and YOLOv9 models uses the mean average precision to assess how well the model detects and segments the multi-class rice seed. It's also an evaluation metric used in this study[21] based on an improved YOLOv8n model for detecting a germinated maize seed. The result in Figure 16 from an IoU of 0. 70, clearly indicates that YOLOv9c-seg is more precise in terms of locating the rice seed and its boundary edges in both the validation and test sets. Both models demonstrate good generalization, with a stable validation and test gap. Therefore, they can detect and segment rice seeds in unseen data, which applies to real-time deployment. Figure 17 shows a higher result of mAP@50 from both models under validation and test sets, which is less complicated than mAP@50-95, and is more precise on the boundary edges of a rice seed. Results such as this are normal due to segmentation complexity and are also an important gauge in measuring the models' generalization on small seeds such as rice, whether overlapping or partially occluded.

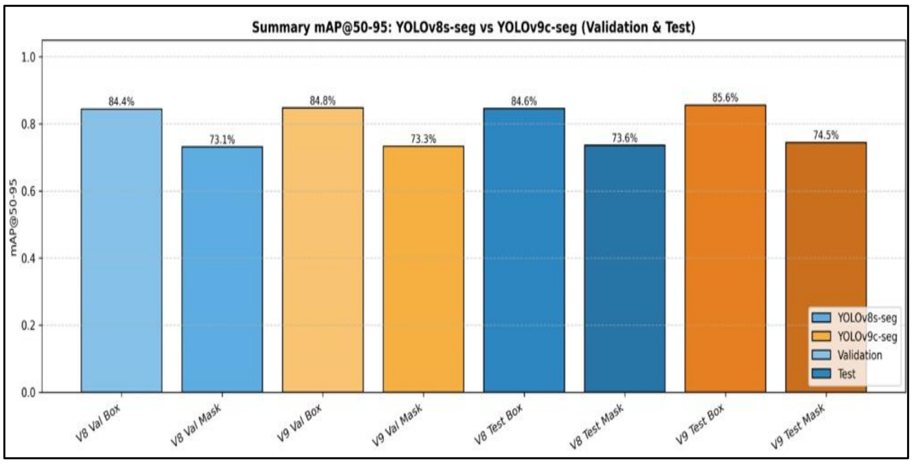

Figure 16. Summary of Mean Average Precision@50-95 (Comparison of YOLOv8s-Seg and YOLOv9c-Seg Performance Metrics on Its Validation and Test Set)

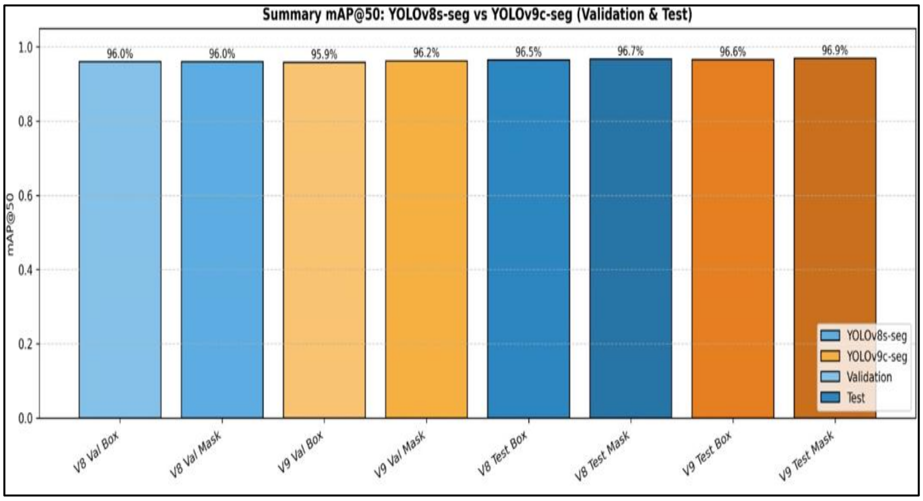

Figure 17. Mean Average Precision@50 (Comparison of YOLOv8s-Seg and YOLOv9c-Seg Performance Metrics on Its Validation and Test Set)

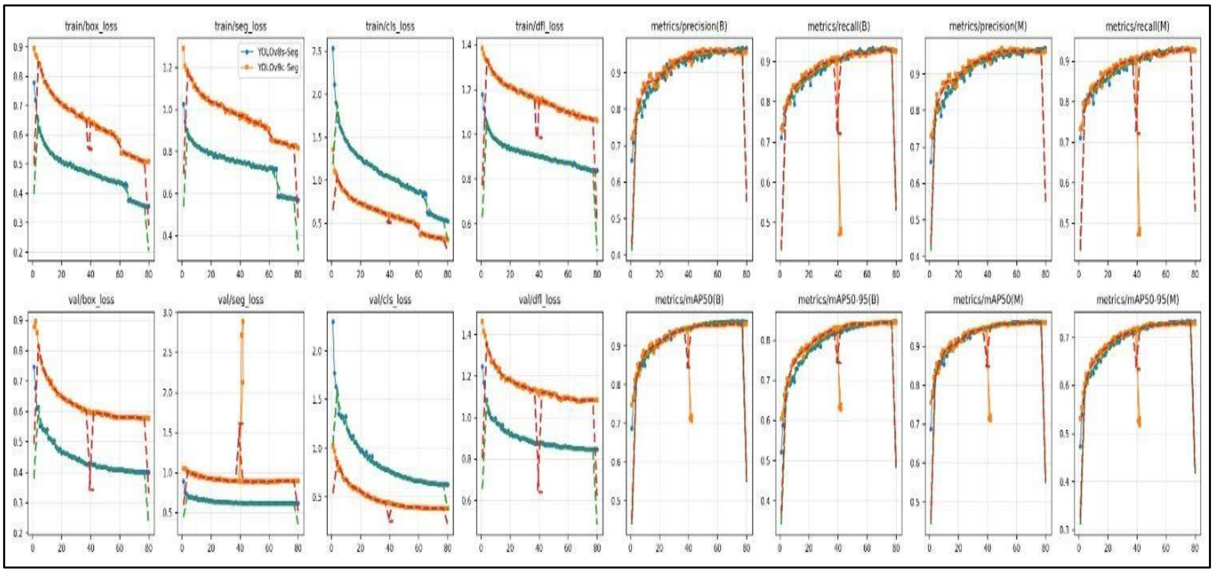

Figure 18. Training History of both Models (Comparison of YOLOv8s-Seg and YOLOv9c-Seg Performance Metrics on Traffic Loss by Training Configuration)

As observed in Figure 18, datasets were trained for 80 epochs, where the training and validation loss are described through the x-axis for the epoch value and the y-axis for the traffic loss value. The training loss curves of both models decrease steadily, resulting in a smooth curve indicating stable optimization; however, in terms of convergence speed, YOLOv9c-seg is faster to reach the final loss value. The validation curve shows a similar downward trend to the training loss; however, a minimal spike is observed at epoch 40, but it recovers quickly due to the presence of datasets where segmentation masks are more sensitive than box detection. In the precision and recall metrics, the x-axis representsthe epoch value, and the y-axis shows the metric value. The precision and recall curves rapidly increase, and both models are very close at high values greater than 90%, with YOLOv9c-seg slightly elevated. Its mAP curves rise quickly and stabilize early for both models, indicating high values that demonstrate strong detection capability. The training history of both models clearly shows no overfitting based on their learning progress.

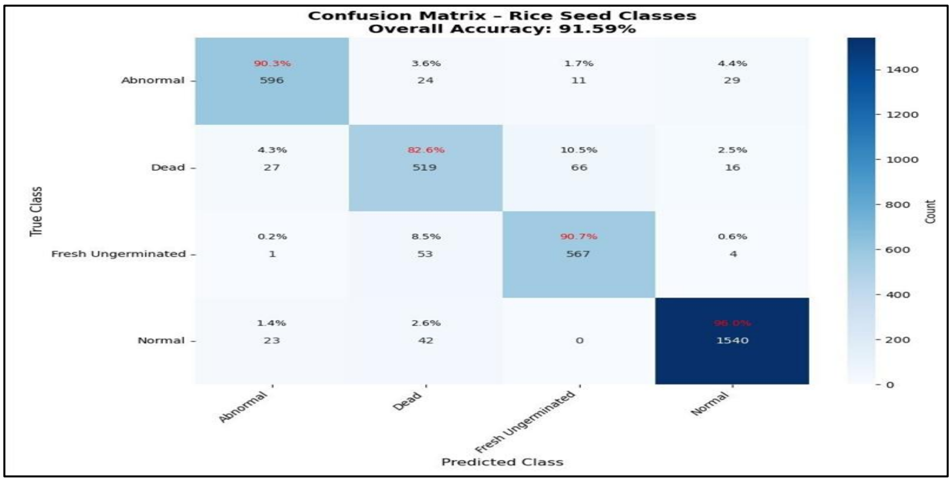

Figure 19. Confusion Matrix of YOLOv8s-Seg

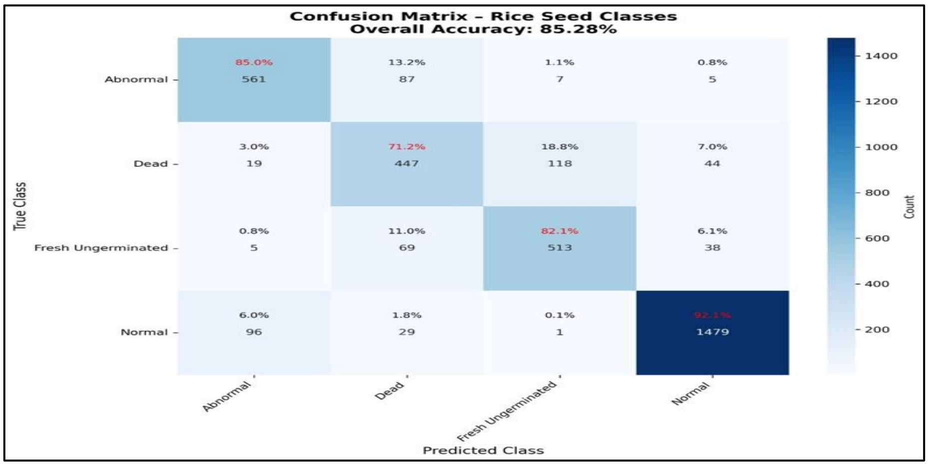

Figure 20. Confusion Matrix of YOLOv9c-Seg

As shown in Figures 19 and 20, the predicted class of a rice seed compared to its true label or ground truth based on the validation dataset expresses a higher accuracy of 91.59% for YOLOv8s-Seg over YOLOV9c-Seg, 85.28%. The frequently misclassified pair detected by both models was the dead and fresh ungerminated classes due to their quite similar visual and morphological structures, specifically in color and texture, where YOLOv9C-Seg bears more false negatives. Having this result may lead to missed objects learned by the model. However, the behavioral progress of both models' detection accuracy is favorable for the multi-class rice seed, as the majority of them were correctly predicted from their ground truth.

4.2 YOLOV8s-Seg and YOLOV9c-Seg Model Performance Metrics (Mixed Dataset)

The performance of two models on a mixed dataset (individual, partially overlapping, and occluded images) was evaluated to determine whether their detection and segmentation results closely match those from the individual dataset. Figure 21 shows the dataset to be used for training, which consists of 11,992 labeled images containing 15,478 instances. A class weighting technique was applied to avoid class imbalance and to form a true result from the ground truth, achieving robust performance.

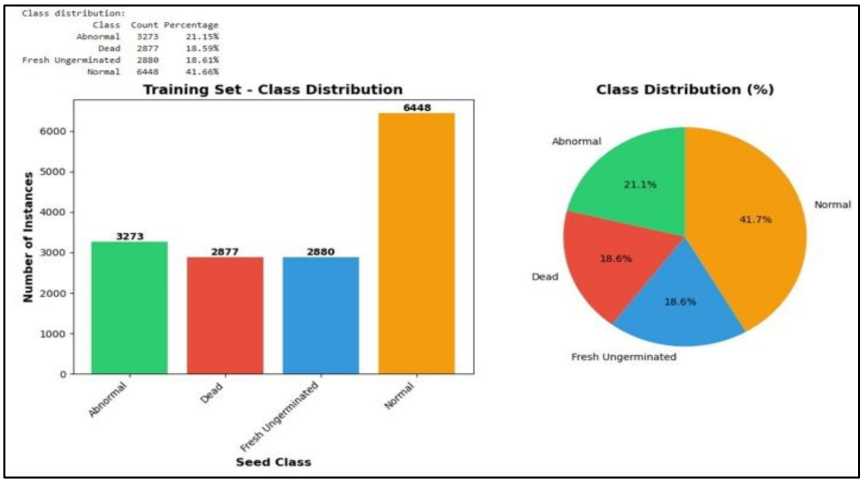

Figure 21. Training Dataset Class Distribution (Mixed Dataset)

The recall summary shown in Figure 22 explains the actual detection of both models in a mixed dataset. As observed, the recall value on the validation and test sets in detection (box) and segmentation (mask) is lower by about 6-8% compared to the recall summary in Figure 14, due to an occluded image introducing harder detection conditions. Even with this result, both models identify most seeds in an unseen area, even if the images are complex. The detection was made easier due to layers of these models extracting rice seed features accordingly. YOLOv9c-Seg slightly dominates its validation and test sets in Figure 23, detecting multi-class seeds labeled as normal, abnormal, fresh ungerminated, and dead over YOLOv8s-Seg. The result still favors the performance of both models in the detection and segmentation of these rice seeds, which are exactly classified.

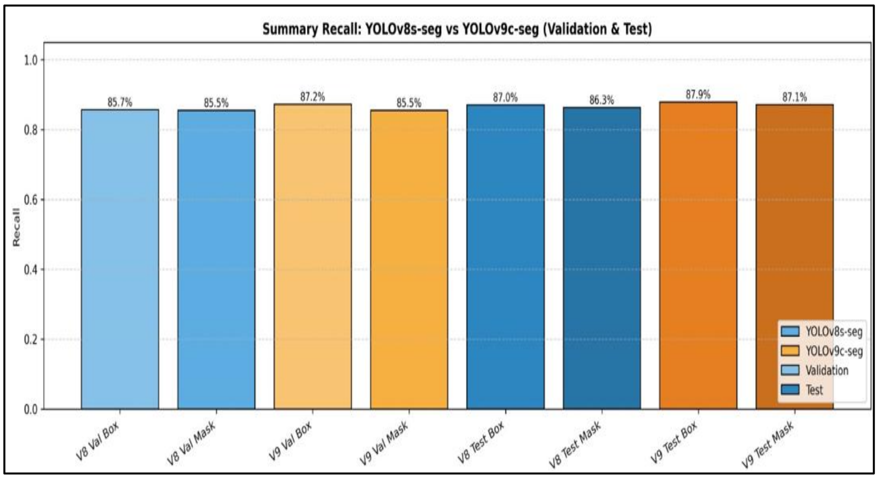

Figure 22. Recall Summary (Comparison of YOLOv8s-Seg and YOLOv9c-Seg Performance Metrics on Its Validation and Test Set)

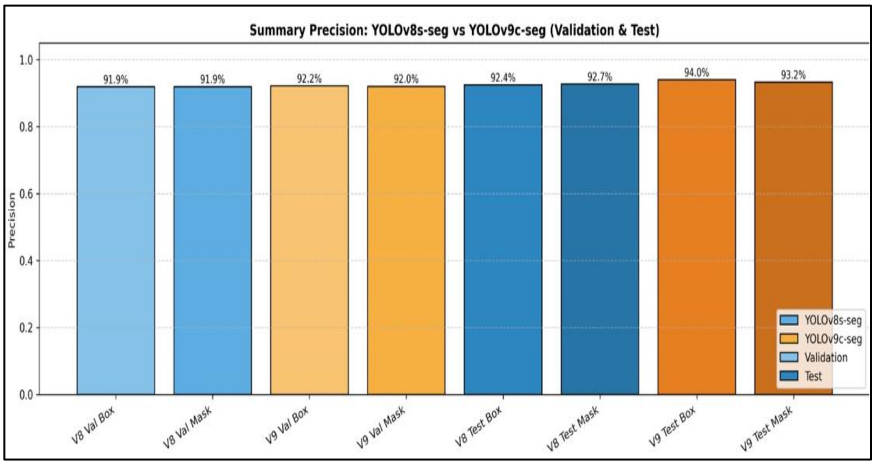

Figure 23. Precision Summary (Comparison of YOLOv8s-Seg and YOLOv9c-Seg Performance Metrics on Its Validation and Test Set)

The mean average precision (mAP) of both models at a stricter IoU of 65% slightly drops from Figure 24, about 6-7% compared to the result from an individual image dataset in Figure 16. The detection makes it complex, specifically for mask segmentation of partially occluded or overlapping images, as it requires pixel-level accuracy. The model also detects incomplete masks, merging, and boundary leakage. YOLOv8s-Seg slightly outperforms the other model in detecting, segmenting, and classifying rice seeds in a mixed dataset. The model's learning capability from the ground truth at a stricter IoU was evident when comparing the raw and detected images seen in Figure 29.

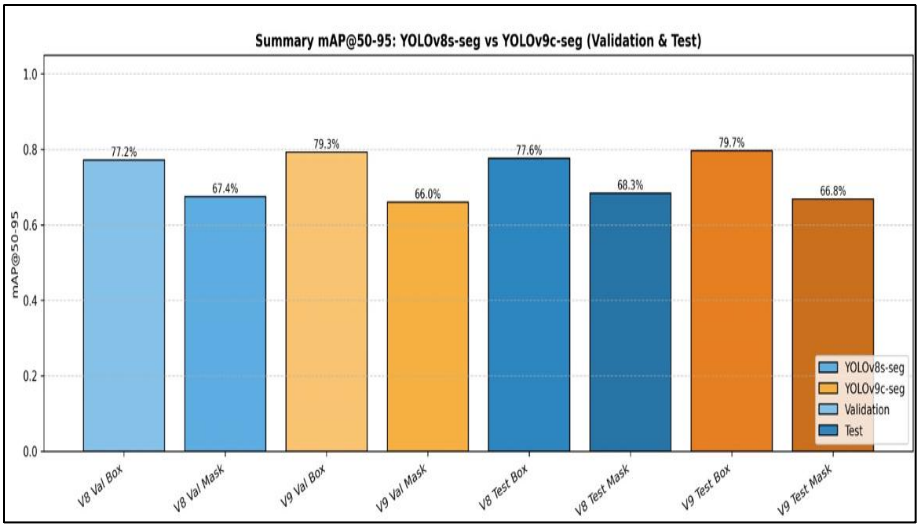

Figure 24. Summary of Mean Average Precision@50-95 (Comparison of YOLOv8s-Seg and YOLOv9c-Seg Performance Metrics on Its Validation and Test Set)

The mAP@50 in Figure 25 displays a higher result based on models learning from a mixed dataset of its validation and test sets. mAP@50 tolerates partial overlap, unlike mAP@50-95, which penalizes boundary inaccuracies caused by occlusion conditions. This observed discrepancy aligns with the known fact that precise localization under occlusion may cause performance degradation at high IoU thresholds. Authors [22] introduce a real-time mask generator that significantly illustrates that mask quality is distinct from box detection quality, and the model used generates a higher mAP@50 even on occluded objects, but experiences a significant drop at mAP@75.

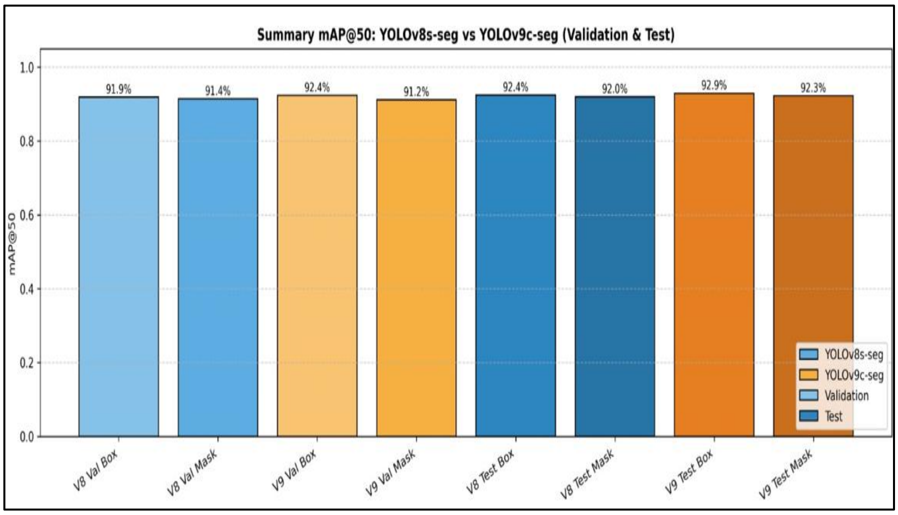

Figure 25. Mean Average Precision@50 (Comparison of YOLOv8s-Seg and YOLOv9c-Seg Performance Metrics on Its Validation and Test Set)

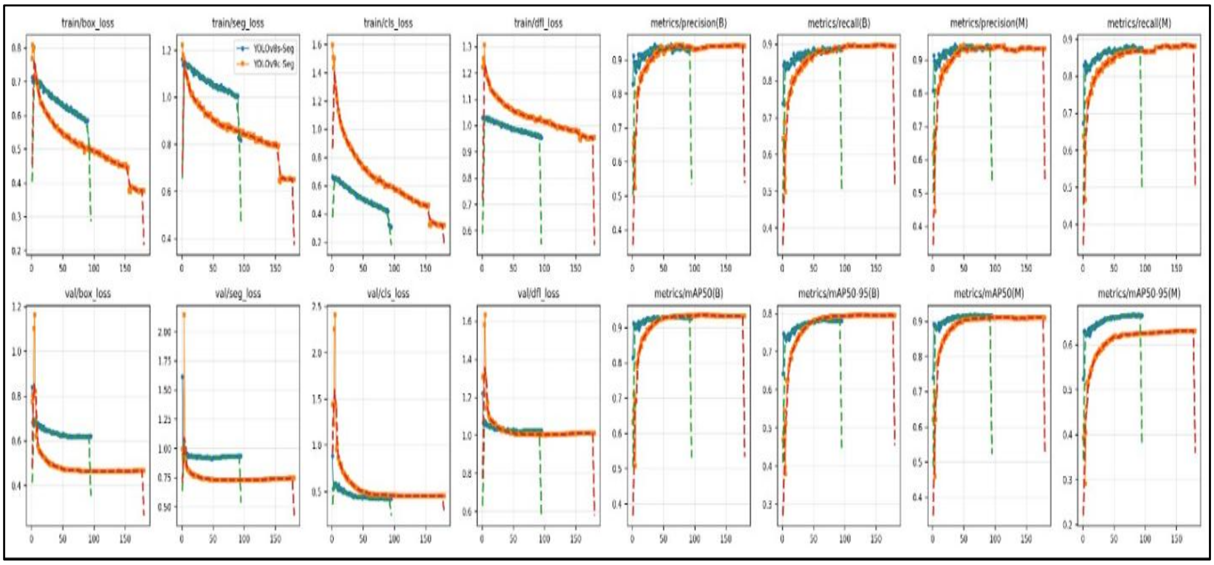

Figure 26. Training History of both Models (Comparison of YOLOv8s-Seg and YOLOv9c-Seg Performance Metrics on Traffic Loss by Training Configuration)

YOLOv8s-Seg and YOLOv9c-Seg were trained for 100 and 180 epochs, respectively, as shown in Figure 26, where their training and validation loss are described on the x-axis for the epoch value and the y-axis for the traffic loss value. The longer epochs used ensure precise boundary and box detection, which works well for images with partial overlap, individuals, or occlusions. From the loss metrics, both models show a declining trend in their learning capability, effectively substantiating the ability to detect, segment, and classify the presence of rice seeds based on their morphological characteristics. In terms of precision and recall metrics in Figure 26, the x-axis represents the epoch value, and the y-axis shows the metric value. Both models achieve acceptable scores above 90% on their mAP@50, effectively learning rice seed objects, despite imperfect pixel quality, which were recognized through mAP@50-95, is the segmentation metric degrades.

Figures 27 and 28 show the confusion matrix of both models, where the accuracy of YOLOv8s-Seg slightly outperforms that of YOLOv9c-Seg; however, these models classify most of the rice seeds. In this case, they efficiently learn classes as normal, abnormal, fresh ungerminated, and dead seeds. Both matrices show the models' performance in detecting true positives of an object, where the majority of the rice seeds were identified and classified accordingly; therefore, both models learn that the small objects are labeled accordingly, with few false positives and false negatives. It was observed that the models trained in a mixed dataset perform significantly better in predicting fresh ungerminated seeds from their ground truth, where a few misclassified dead seeds were detected. For abnormal rice seeds, the models are slightly confused and classify them as normal rice seeds due to the presence of their germination structures, such as roots and shoots. Missed objects are likely present because of false negatives detected by both models.

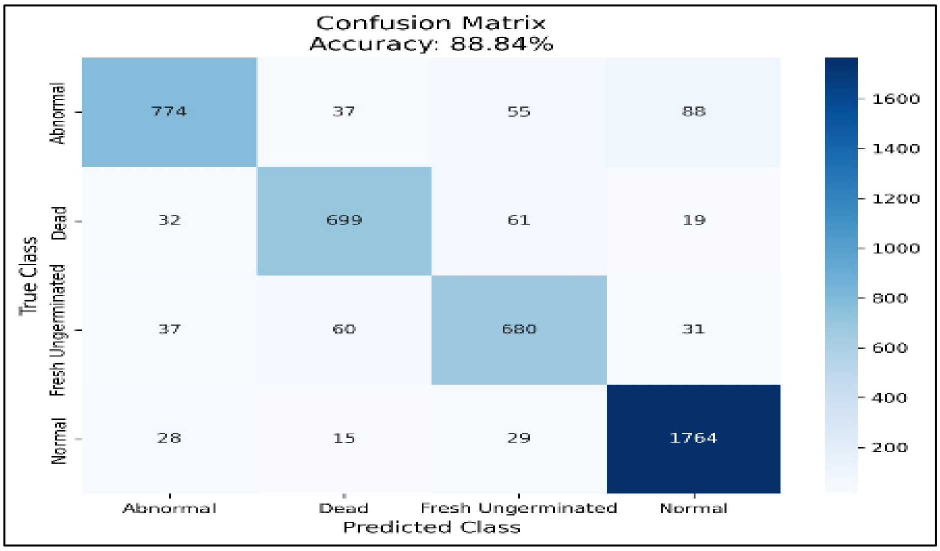

Figure 27. Confusion Matrix of YOLOv8s-Seg

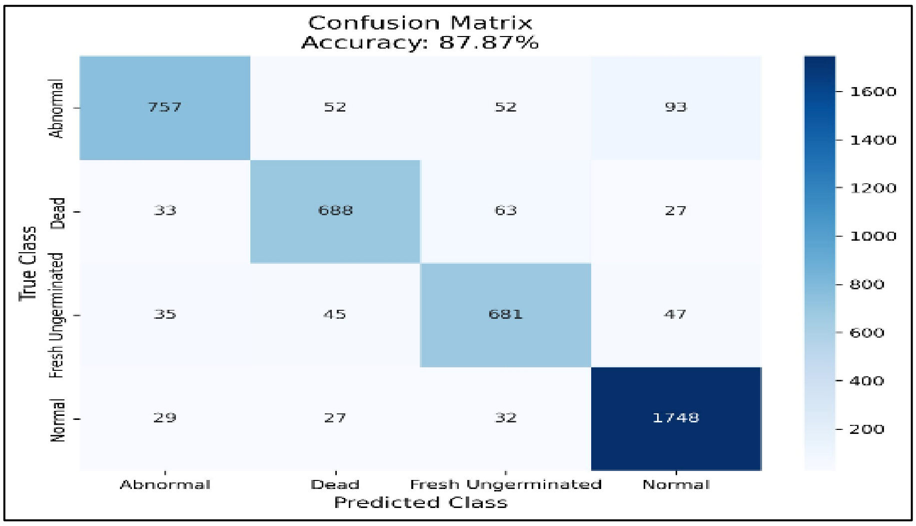

Figure 28. Confusion Matrix of YOLOv9c-Seg

4.3 Post Processing (Integration of Segment Anything Model to YOLOv8s-Seg and YOLOv9c-Seg)

Figure 29 displays the outcome of integrating YOLO's instance segmentation with the Segment Anything Model (SAM) in the image. It correctly defines boundaries on an individual seed; however, when YOLOv8s-Seg detects and segments multiple images in one frame, some classes are missed, and only a few are detected, which means the model is not completely learning from it.

Figure 30 uses YOLOv9c-Seg for the individual dataset and finds an improvement in detection and segmentation. The only difference it had from Figure 29 was its multiple class seed, where the instance segmentation of YOLO detects 6 seeds. As it progresses with the labeler refinement SAM, it detects 7 seeds, but some seeds were missed and misclassified. Figure 31 uses YOLOv8s-Seg to train a mixed dataset, and the outcome of its detection and refined segmentation correctly identifies the rice seed objects according to their classes, even under occluded conditions, which has significantly improved accuracy.

4.3.1 For Individual Dataset (YOLOv8s-Seg)

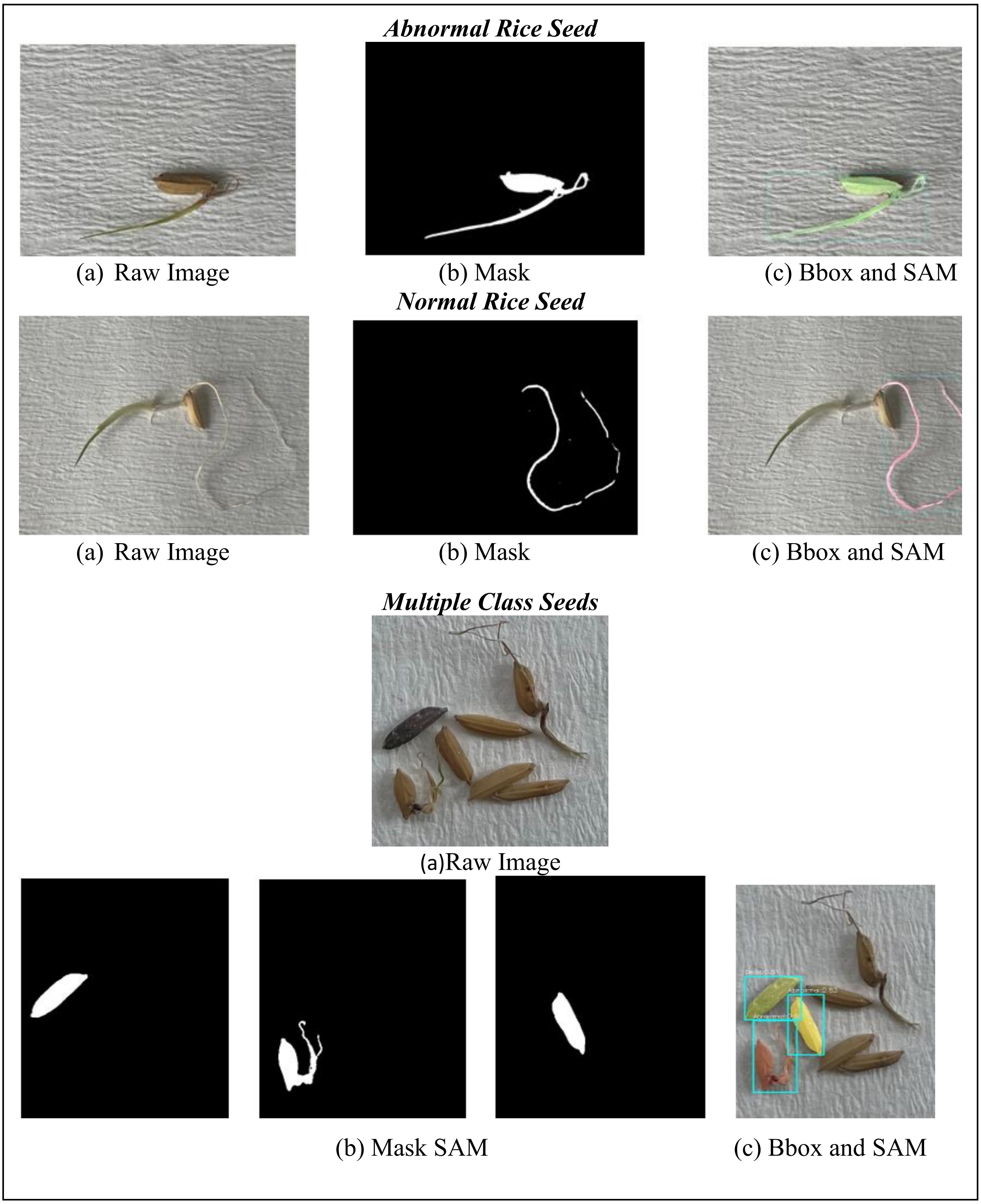

Figure 29. Inference Output using YOLOv8s-Seg and SAM (Individual Dataset)

4.3.2 For Individual Dataset (YOLOv9c-Seg)

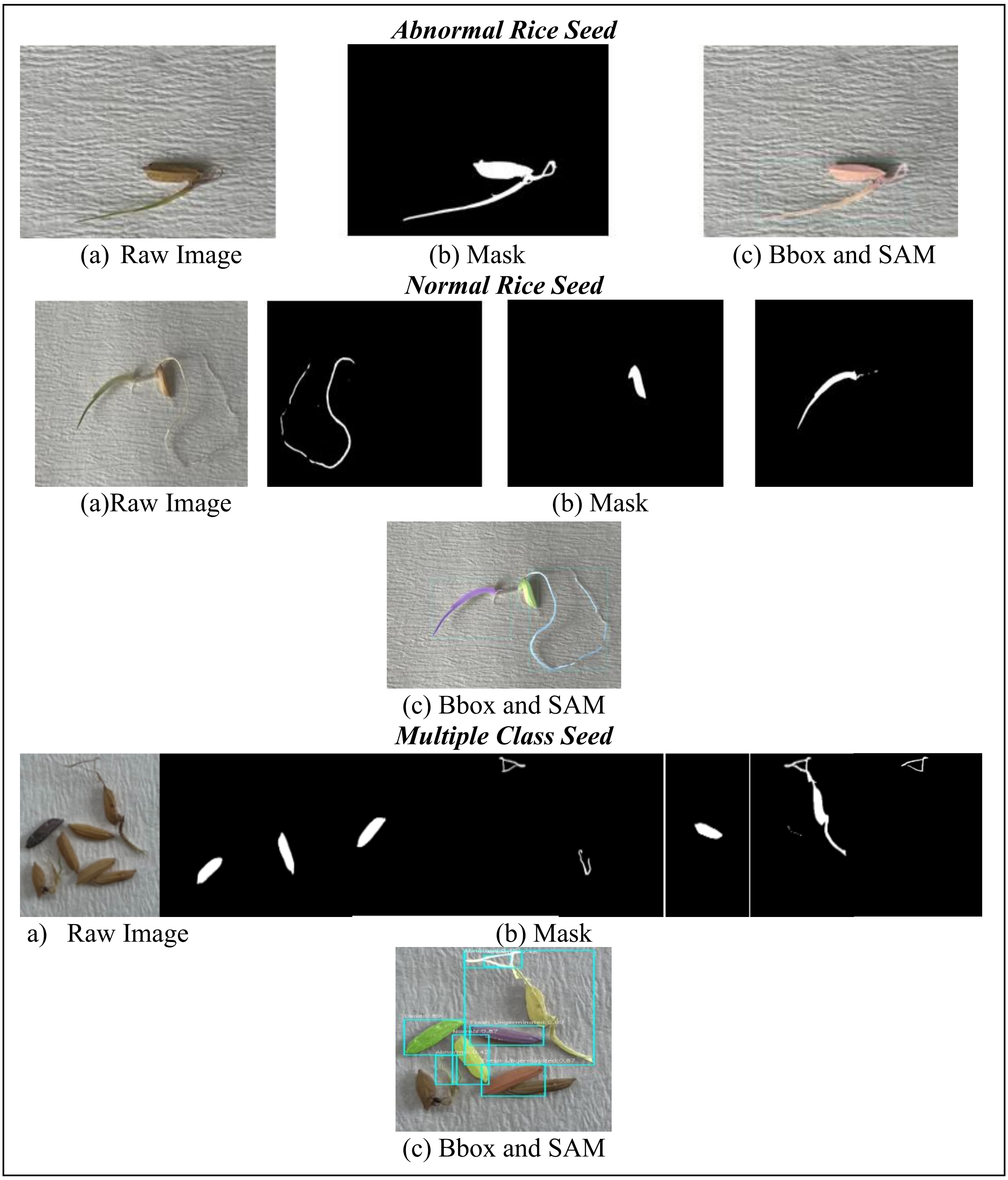

Figure 30. Inference Output using YOLOv9c-Seg and SAM (Individual Dataset)

4.3.3 For Mixed Dataset (YOLOv8s-Seg)

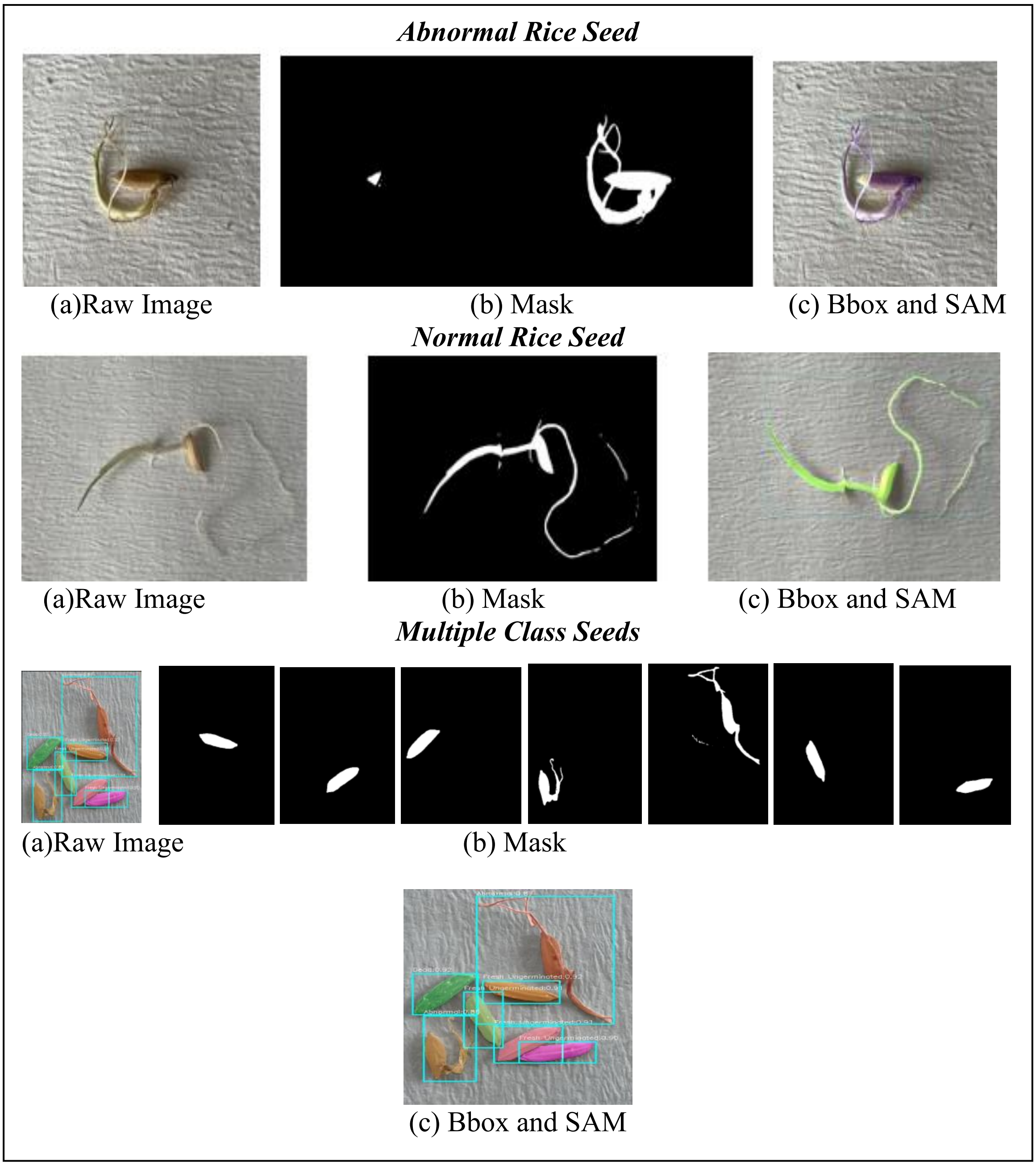

Figure 31. Inference Output using YOLOv8s-Seg and SAM (Mixed Dataset)

The YOLOv9c-Seg model was used to train a mixed dataset that contains individual and multi-class rice seeds, as shown in Figure 32. The output of its detection and segmentation, from individual rice seeds to multiple class seeds has been accurately identified based on its refined mask and box detection. However, duplicate detections are appearing since labelling or annotating the rice seeds is fairly challenging, especially in partially overlapping and occluded conditions, which sometimes leads to confusion in the model's learning.

4.3.4 For Mixed Dataset (YOLOv9c-Seg)

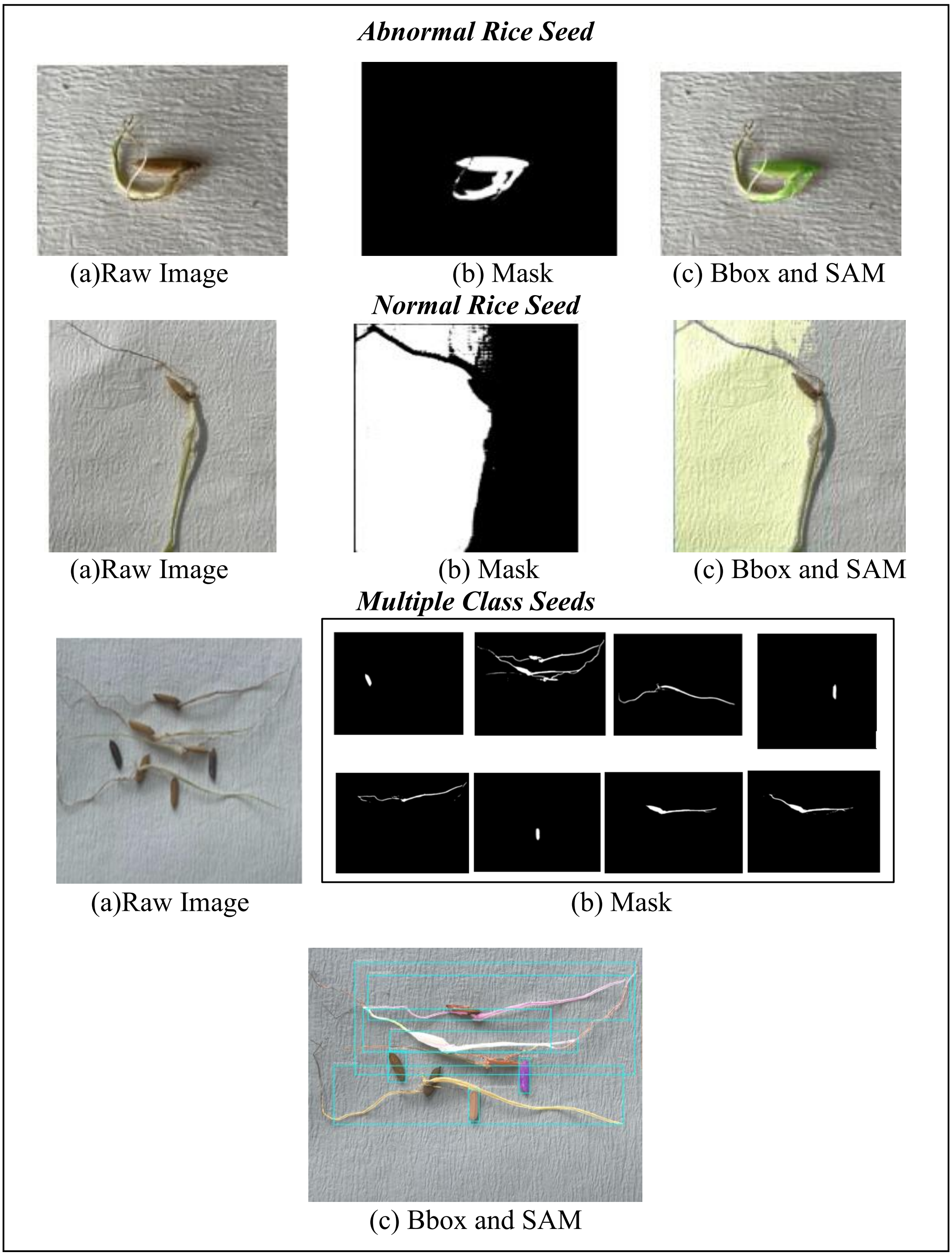

Figure 32. Inference Output using YOLOv9c-Seg and SAM (Mixed Dataset)

The model's capability to accurately detect and segment in the later training is dependent on how it was labeled or annotated, especially in partially overlapping and occluded situations. However, polygon annotations using Roboflow can mitigate challenges in manual annotation by following its predefined annotation guidelines, ensuring that seed contours are traced to form the rice seed structure from various classes, thus avoiding boundary inaccuracies.

This study compared the performance of two models, YOLOv8s-Seg and YOLOv9c-Seg, in their ability to detect and segment individual and multi-class germinated rice seeds as normal, abnormal, fresh ungerminated, and dead seeds. In summary, the model will not completely learn the partially overlapping or occluded conditions if the dataset to be processed consists solely of individual seeds. It is better for the model to learn in a multi-class environment so that detection and segmentation can be defined and labeled accordingly, resulting in no missed objects.

The inference output in this study was processed on pre-captured images with the integration of SAM for coarse mask refinement from YOLO's instance mask segmentation, which mitigates boundary-related errors, especially in partially overlapping or occluded conditions. This integration helps to improve the segmentation process of an image, specifically for smaller objects such as rice seeds. Both models, YOLOv8s-Seg and YOLOv9c-Seg, were suitable for web and mobile deployment, with YOLOv8s-Seg being appropriate for mobile applications due to its computational complexity designed for small objects, and it may yield better results in detecting, segmenting, and classifying germinated rice seeds as normal, abnormal, fresh ungerminated, and dead seeds.

5. Conclusion and Future Works

This paper presents the results of CNN-based technologies, in this case, YOLOv8s-Seg and YOLOv9c-Seg, to perform multi-class analysis of germinated rice seeds by detecting, segmenting, and classifying the radicle, as well as color, surface texture, and edge boundary features. YOLOv8s-Seg is light enough to be deployed in a mobile application. YOLOv9c-Seg is more accurate but heavier on computation, making it more difficult to deploy in real-time scenarios. The performance of both models was good, but misclassified instances occurred between visually similar seed classes. In summary, the use of CNNs will help in monitoring seed quality, and manual identification methods may be replaced. There is potential for a mobile-based application in precision agriculture.

References

[1] Debangshi, Udit, Aishik Sadhukhan, Darshan Dutta, and Sougata Roy. "Application of Smart Farming Technologies in Sustainable Agriculture Development: A comprehensive Review on Present Status and Future Advancements." International Journal of Environment and Climate Change 13, no. 11 (2023): 3689-3704.

[2] Virk, Ahmad Latif, Mehmood Ali Noor, Sajid Fiaz, Saddam Hussain, Hafiz Athar Hussain, Muzammal Rehman, Muhammad Ahsan, and Wei Ma. "Smart Farming: An Overview." Smart village technology: concepts and developments (2020): 191-201.

[3] Hidayat, Sidiq Syamsul, Dwi Rahmawati, Muhamad Cahyo Ardi Prabowo, Liliek Triyono, and Farika Tono Putri. "Determining the Rice Seeds Quality Using Convolutional Neural Network." JOIV: International Journal on Informatics Visualization 7, no. 2 (2023): 527-534.

[4] Durai, S., C. Mahesh, T. Sujithra, and C. Shyamalakumari. "Germination Prediction System for Rice seed using CNN Pre-Trained Models." In 2022 International Conference

on Advances in Computing, Communication and Applied Informatics (ACCAI), IEEE, 2022, 1-9.

[5] Li, Huikang, Longbao Liu, Qi Li, Juan Liao, Lu Liu, Yujun Zhang, Qixing Tang, Yuan Rao, and Yanwei Gao. "RSG-YOLOV8: Detection of Rice Seed Germination Rate Based on Enhanced YOLOv8 and Multi-Scale Attention Feature Fusion." PLoS One 19, no. 11 (2024): e0306436.

[6] Yao, Qiong, Xiaoming Zheng, Guomin Zhou, and Jianhua Zhang. "SGR-YOLO: A Method for Detecting Seed Germination Rate in Wild Rice." Frontiers in Plant Science 14 (2024): 1305081.

[7] Da Silva, Daniel de Amaral, Emannuel Diego Gonçalves De Freitas, Haynna Fernandes Abud, and Danielo G. Gomes. "Applying YOLOv8 and X-ray Morphology Analysis to Assess the Vigor of Brachiaria brizantha cv. Xaraés Seeds." AgriEngineering 6, no. 2 (2024): 869-880.

[8] Wu, Yijie, Zhengjun Li, Haoyu Jiang, Qianyun Li, Jinxin Qiao, Feng Pan, Xiuqing Fu, and Biao Guo. "YOLOv8-SegANDcal: Segmentation, Extraction, and Calculation of Soybean Radicle Features." Frontiers in plant science 15 (2024): 1425100.

[9] Brutas, Mariel John B., Arthur L. Fajardo, Erwin P. Quilloy, Luther John R. Manuel, and Adrian A. Borja. "Enhancing Seed Germination Test Classification for Pole Sitao (Vigna unguiculata (L.) Walp.) using SSD MobileNet and Faster R-CNN models." Applied Sciences 14, no. 13 (2024): 5572.

[10] Sun, Wenbin, Meihan Xu, Kang Xu, Dongquan Chen, Jianhua Wang, Ranbing Yang, Quanquan Chen, and Songmei Yang. "CSGD-YOLO: A Corn Seed Germination Status Detection Model Based on YOLOv8n." Agronomy 15, no. 1 (2025): 128.

[11] Sha, Mo, Xiuqing Fu, Ruxiao Bai, Zhibo Zhong, Haoyu Jiang, Fei Li, and Siyu Yang. "YOLOv8-Licorice: A Lightweight Salt-Resistance Detection Method for Licorice Based on Seed Germination State." Frontiers in Plant Science 15 (2024): 1474321.

[12] Yue, Xiang, Kai Qi, Xinyi Na, Yang Zhang, Yanhua Liu, and Cuihong Liu. "Improved YOLOv8-Seg Network for Instance Segmentation of Healthy and Diseased Tomato Plants in the Growth Stage." Agriculture 13, no. 8 (2023): 1643.

[13] Jiang, Haoyu, Fei Hu, Xiuqing Fu, Cairong Chen, Chen Wang, Luxu Tian, and Yuran Shi. "YOLOv8-Peas: A Lightweight Drought Tolerance Method for Peas Based on Seed Germination Vigor." Frontiers in Plant Science 14 (2023): 1257947.

[14] Sapkota, Ranjan, Dawood Ahmed, and Manoj Karkee. "Comparing YOLOv8 and Mask R-CNN for Instance Segmentation in Complex Orchard Environments." Artificial Intelligence in Agriculture 13 (2024): 84-99.

[15] Khalid, Mahnoor, Muhammad Shahzad Sarfraz, Uzair Iqbal, Muhammad Umar Aftab, Gniewko Niedbała, and Hafiz Tayyab Rauf. "Real-Time Plant Health Detection Using Deep Convolutional Neural Networks." Agriculture 13, no. 2 (2023): 510.

[16] Almalky, Abeer M., and Khaled R. Ahmed. "Deep Learning for Detecting and Classifying the Growth Stages of Consolida Regalis Weeds on Fields." Agronomy 13, no. 3 (2023): 934.

[17] Toda, Yosuke, Fumio Okura, Jun Ito, Satoshi Okada, Toshinori Kinoshita, Hiroyuki Tsuji, and Daisuke Saisho. "Training Instance Segmentation Neural Network with Synthetic Datasets for Crop Seed Phenotyping." Communications biology 3, no. 1 (2020): 173.

[18] Yu, Helong, Jiayao Zhao, Chunguang Bi, Jing Chen, and Ming Zhao. "Study on the Germination Rate of Maize Seeds Based on Improved YOLOv8n Model." Frontiers in Plant Science 16 (2025): 1555440.

[19] Liang, Baihe, Liangchen Hu, Guangxing Liu, Peng Hu, Shaosheng Xu, and Biao Jie. "YOLOv9-GSSA Model for Efficient Soybean Seedlings and Weeds Detection." Smart Agricultural Technology (2025): 101134.

[20] Zhu, Guanquan, Zihang Luo, Minyi Ye, Zewen Xie, Xiaolin Luo, Hanhong Hu, Yinglin Wang, Zhenyu Ke, Jiaguo Jiang, and Wenlong Wang. "Instance Segmentation of Sugar Apple (Annona squamosa) in Natural Orchard Scenes Using an Improved YOLOv9-Seg Model." Agriculture; Basel 15, no. 12 (2025).

[21] Wang, Qiaolong, Dongshun Chen, Wenfei Feng, Liang Sun, and Gaohong Yu. "Enhancing Instance Segmentation in Agriculture: An Optimized YOLOv8 Solution." Sensors 25, no. 17 (2025): 5506.

[22] Bolya, Daniel, Chong Zhou, Fanyi Xiao, and Yong Jae Lee. "Yolact: Real-Time Instance Segmentation." In Proceedings of the IEEE/CVF international conference on computer vision, 2019, 9157-9166.