

## Article

# Detection of Aging Maize Seed Vigor and Calculation of Germ Growth Speed Using an Improved YOLOv8-Seg Network

Helong Yu <sup>1,2</sup>, Xi Ling <sup>1</sup>, Zhenyang Chen <sup>2</sup>, Chunguang Bi <sup>1,\*</sup> and Wanwu Zhang <sup>3,\*</sup>

<sup>1</sup> College of Information Technology, Jilin Agricultural University, Changchun 130118, China; yuhelong@aliyun.com (H.Y.); 20221154@mails.jlau.edu.cn (X.L.)

<sup>2</sup> Smart Agriculture Research Institute, Jilin Agricultural University, Changchun 130118, China

<sup>3</sup> Jilin Rural Economy Information Center, Changchun 130000, China

\* Correspondence: chunguang@jlau.edu.cn (C.B.); zhangwanwu2024@163.com (W.Z.)

**Abstract:** Crop yields are influenced by various factors, including seed quality and environmental conditions. Detecting seed vigor is a critical task for seed researchers, as it plays a vital role in seed quality assessment. Traditionally, this evaluation was performed manually, which is time-consuming and labor-intensive. To address this limitation, this study integrates the ConvUpDownModule (a customized convolutional module), C2f-DSCConv (C2f module with Integrated Dynamic Snake Convolution), and T-SPPF (the SPPF module integrated with the transformer multi-head attention mechanism) into the VT-YOLOv8-Seg network (the improved YOLOv8-Seg Network), an enhancement of the YOLOv8-Seg architecture. The ConvUpDownModule reduces the computational complexity and model parameters. The C2f-DSCConv leverages flexible convolutional kernels to enhance the accuracy of maize germ edge segmentation. The T-SPPF integrates global information to improve multi-scale segmentation performance. The proposed model is designed for detecting and segmenting maize seeds and germs, facilitating seed germination detection and germination speed computation. In detection tasks, the VT-YOLOv8-Seg model achieved 97.3% accuracy, 97.9% recall, and 98.5% mAP50, while in segmentation tasks, it demonstrated 97.2% accuracy, 97.7% recall, and 98.2% mAP50. Comparative experiments with Mask R-CNN, YOLOv5-Seg, and YOLOv7-Seg further validated the superior performance of our model in both detection and segmentation. Additionally, the impact of seed aging on maize seed growth and development was investigated through artificial aging studies. Key metrics such as germination rate and germ growth speed, both closely associated with germination vigor, were analyzed, demonstrating the effectiveness of the proposed approach for seed vigor assessment.

Academic Editors: Marcus Randall and Gniewko Niedbala

Received: 13 December 2024

Revised: 26 January 2025

Accepted: 31 January 2025

Published: 1 February 2025

**Citation:** Yu, H.; Ling, X.; Chen, Z.; Bi, C.; Zhang, W. Detection of Aging Maize Seed Vigor and Calculation of Germ Growth Speed Using an Improved YOLOv8-Seg Network. *Agriculture* **2025**, *15*, 325. <https://doi.org/10.3390/agriculture15030325>

**Copyright:** © 2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** seed vigor detection; germination rate; YOLOv8-Seg; artificial seed aging; computer vision in agriculture

## 1. Introduction

As far as global production volume goes, maize (*Zea mays* L.) is the most important staple crop [1]. There are over 205 million hectares of maize grown worldwide, and about 1.2 billion MT of maize are produced overall [2]. Its significance underscores the critical need for evaluating seed quality, particularly seed vigor, which is essential for the seed industry [3]. Seed vigor plays a pivotal role in crop production, as crop yield and resource-use efficiency depend heavily on the successful establishment of plants in the field.

Specifically, seed vigor determines the ability of seeds to germinate and establish uniformly and robustly under diverse environmental conditions [4]. This makes it a vital metric for assessing the physiological potential of seeds at various production stages. Moreover, seed vigor evaluation aids in selecting high-quality seed lots to meet consumer demand, further highlighting its relevance. It is a critical tool for assessing the physiological potential of seeds at various stages of production. Furthermore, it informs the selection of high-quality seed lots to meet consumer demand. As a key indicator of seed quality which directly influences agricultural yields, evaluating seed germination remains a fundamental task for seed researchers [5].

Conventional techniques for evaluating seed quality, especially seed viability testing, mostly depend on labor-intensive and time-consuming manual measurements. The traditional germination tests are prone to subjective errors and frequently restrict the scalability and efficiency of the evaluations, which results in poor reproducibility and inconsistent statistical results between evaluators. An objective, repeatable, quick, and financially feasible measurement technique is therefore desperately needed. The majority of seed testing methods set by the International Seed Testing Association (ISTA) are assessed manually using a variety of standardized processes that are adapted for certain crops [6]. Therefore, there is a pressing need for quicker and more effective ways to identify when seeds are germinating.

Modern image analysis technologies have been applied to seed detection due to their ability to automate the process and provide quantitative measurements with minimal error [7]. Systems such as GERMINATOR and the Seed Vigor Imaging System (SVIS) have been developed to evaluate seed germination and viability [8,9]. GERMINATOR analyzes color contrast between the radicle and testa to assess germination in species like Brassica and Arabidopsis, while SVIS determines seed viability by measuring seed length from RGB pixel data. Despite their utility, these methods still require manual input or specialized equipment, limiting their generalizability and efficiency.

The agricultural industry has made significant use of machine vision technology, which has seen undergone substantial advancements [10–13]. Researchers have developed techniques that use machine vision technology to check the quality of seeds [14–16]. To forecast seed germination, Reference [17] used a combination of image mapping and linear discriminant analysis models. Reference [18] utilized the K-nearest neighbors (KNN) model to enhance the scoring accuracy of mangosteen seeds. To evaluate the germination status of rice seeds in Thailand, Reference [19] created the Rice Seed Germination Evaluation System (RSGES), which is based on artificial neural network classifiers. To categorize the viability of lentil seeds, Reference [20] used interactive machine learning and seedling image analysis approaches. However, the manual feature extraction used by classic machine learning classifiers leads to slower training periods and worse accuracy.

Significant progress has been made in deep learning in recent years, with widespread applications in agriculture. This includes its use in germination trays and Petri dishes for tasks such as seed detection, identification, and germination analysis [21–23]. References [24,25] integrated deep learning with machine learning approaches to develop large-scale germination scoring systems. These systems leverage UNet, Gaussian mixture models, and other techniques to identify germination-related features across diverse crop species. Additionally, convolutional neural network (CNN)-based detectors such as Mask R-CNN and Faster R-CNN have been used to extract regions of interest for seed germination detection [26,27]. However, their reliance on proprietary datasets and high computational requirements limits their accessibility.

YOLO (You Only Look Once) series algorithms have emerged as a robust alternative, offering a lightweight architecture and fast detection capabilities. Researchers have utilized YOLO-based models for seed germination detection with remarkable success. For

example, Reference [28] developed the Wheat Seed Viability Assessment System (WSVAS) using YOLOv4, achieving a mean average precision (mAP) of 97.59% and a recall rate of 97.35%. Similarly, Reference [29–31] introduced models like SGR-YOLO and CSGD-YOLO, which further optimized computational efficiency and detection accuracy across various seed types. To detect the germination status of maize seeds at different time intervals, Chen et al. [32] established a Petri dish-based dataset of maize seed germination status at different germination times. Despite these advancements, existing studies primarily focus on germination detection, lacking comprehensive metrics such as seed vigor and germ growth rate evaluation.

To determine the germination rate and speed of maize seeds, this study developed a deep learning model capable of detecting and segmenting seed germination at various aging stages. Three distinct aging experiments were conducted in an aging chamber, and germination experiments were performed using Petri dishes at room temperature. Photographs were captured at one-hour intervals to document the germination process. Since seed germination rate is the most significant factor in the seed vigor evaluation index, germination detection serves as a crucial step. However, the existing research lacks methods that integrate germination rate detection with segmentation, particularly considering the variation in germ size caused by differences in seed vigor during the germination stage. This study addresses these gaps by proposing an innovative approach for seed vigor evaluation using a modified YOLOv8-Seg model. The contributions of this research are as follows:

- (1) Development of the VT-YOLOv8-Seg model: Based on the YOLOv8n model, we introduced several enhancements, including the ConvUpDownModule, C2f-DSCConv, and the T-SPPF. These improvements enable the precise segmentation of maize seed bodies and buds.
- (2) Integration of instance segmentation for germination analysis: The proposed model applies instance segmentation techniques to extract germ areas, enabling accurate calculations of germination speed and providing a comprehensive evaluation of seed quality.
- (3) Application to aging studies: The model was utilized to analyze germination indexes of maize seeds after artificial aging treatments, offering insights into the effects of seed aging on crop growth and development.

We provide a fast, accurate, and scalable method for intelligent seed germination detection, offering a practical solution for evaluating seed vigor in agricultural production. By addressing both germination rate and seed vigor assessment, this study not only contributes to the field of automated seed evaluation, but also serves as a valuable reference for future studies on crop development and seed aging effects.

## 2. Materials and Methods

### 2.1. Data Acquisition

To investigate seed vigor, seeds were classified based on their surface cleanliness and consistency. The selection criteria for seeds included uniformity in size, color, and apparent physical condition to minimize confounding factors during analysis. Damaged, discolored, or irregularly shaped seeds were excluded. The aging durations were set at T0 (control group), T1 (24 h), T2 (48 h), and T3 (72 h). Thirty seeds were selected for each group. Three sets of seeds underwent aging experiments under conditions of 45 °C and humidity levels exceeding 95%, ensuring a high-temperature, high-humidity environment. After aging, the seeds were removed and stored in a cool, well-ventilated indoor environment.

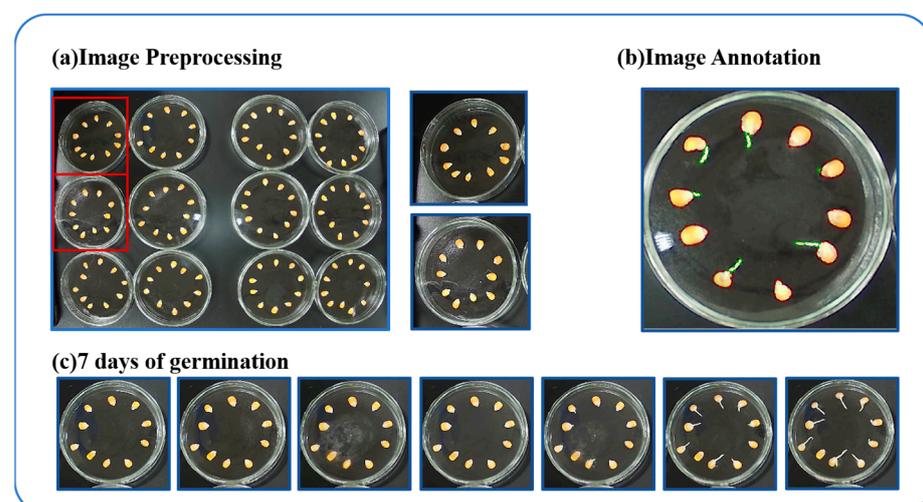
Germination studies were subsequently conducted at room temperature to compare the aged seeds with the original, non-aged samples. Standardized germination tests, conducted according to national standards, were used to assess factors such as germination rate, mean germination time and other relevant metrics. To enhance the visibility of the growing radicle against the background, the seeds were placed in Petri dishes on black cloth. The cloth was moistened with tap water, and the Petri dishes were covered to minimize the effects of water evaporation and seed desiccation. The grid cover may occasionally produce reflections, depending on the camera settings.

The various stages of germination were recorded using digital imaging techniques. One image was captured every hour, resulting in 24 images per day and 168 images over the course of seven days, totaling 2016 images. The study aimed to classify these images based on individual Petri dishes. All images were taken at room temperature in an office setting. Lighting conditions were standardized using a combination of an artificial light source and the built-in light from the camera. Specifically, an office desk lamp with a color temperature of approximately 4000K was paired with the camera's integrated LED light to provide uniform illumination. This lighting arrangement remained consistent throughout the entire shoot, which lasted approximately seven days. The Petri dishes were arranged in a 4x3 grid, with columns designated as T0, T1, T2, and T3. An industrial camera (YT-CAM8008LA-4K/4G) was positioned at a fixed height of 20 cm above the Petri dishes to provide a stable and consistent view for recording the germination process.

## 2.2. Data Pre-Processing

As shown in Figure 1a, the image was segmented into smaller, uniformly sized sections, each containing a single Petri dish. As shown in Figure 1b, bounding boxes were created around each seed using the open-source program LabelMe, enabling seed labeling and polygonal annotations. Each image had a resolution of  $640 \times 640$  pixels and was in JPEG format. The JSON-formatted labeling file was converted into a TXT file containing multiple coordinate points.

Germination is the process by which the radicle emerges from the seed coat. To train and validate the model in this study, 120 seeds were used, resulting in 2016 seed images. Ten seeds were placed in each Petri dish, and observations were systematically conducted every hour for seven days (168 h) after the initial placement. The successful completion of the experiment resulted in 168 images of a single seed captured using this technology. The changes in the seed over the seven-day period are shown in Figure 1c, which includes images selected from the initial observations taken every 24 h.



**Figure 1.** (a) illustrates images showing the seed germination process within Petri dishes. The images were then cropped to focus on a single Petri dish in each instance. (b) An example from the LabelMe dataset is shown. (c) Seed changes were documented over a period of seven days. The first image from each seven-day interval was selected for the presentation of the dataset.

Online data augmentation, which entails real-time augmentation operations of the dataset, is integrated into the YOLOv8 network during the model training phase. During training, online data augmentation techniques such as HSV augmentation, random perspective, MixUp, and mosaic are applied. Since YOLOv8 performs online augmentation to enhance the diversity and generalization of the model's dataset, additional offline dataset augmentation is unnecessary to optimize training efficiency.

A total of 2,016 images of maize seed germination were collected for this study. Subsequently, the entire dataset was divided into training, validation, and test sets with a ratio of 7:2:1.

### 2.3. Network Model Construction

#### 2.3.1. Structure of the YOLOv8-Seg Network

The exceptional accuracy and rapid detection capabilities of the YOLO (You Only Look Once) algorithmic framework distinguish it from other detection techniques. Due to its frequent updates and iterations, the YOLO family of models has gained widespread recognition as a real-time object detection system. It is widely utilized to detect and classify crops, pests, diseases, and other relevant aspects in automated and precision agriculture.

YOLOv8 is a modern object detection algorithm developed by the Ultralytics team, which significantly improves upon YOLOv5 by adopting an anchor-free, one-stage detection framework [32]. Its architecture is divided into four main components: the input module, backbone, neck, and head. The input module integrates several advanced features, such as adaptive image scaling, adaptive anchor calculation, and mosaic data augmentation, enabling dynamic adjustments to input images based on model requirements. Larger models incorporate techniques like MixUp and CopyPaste to enhance data diversity and improve generalization performance.

The backbone utilizes an enhanced ELAN design principle, replacing the C3 structure in YOLOv5 with the more advanced C2f structure. This modification introduces additional residual connections and branching, optimizing feature extraction while maintaining a lightweight design. Adjusting the number of channels based on the model's scale further improves performance.

The neck incorporates path aggregation network (PAN) [33] and feature pyramid network (FPN) [34] designs to effectively integrate features across different scales. This structure enhances the ability to localize objects and capture semantic information, ensuring robust multiscale feature fusion.

The head employs a decoupled structure, separating the classification and localization tasks to enhance detection precision. By adopting an anchor-free approach, it eliminates the need for traditional anchor boxes, significantly improving detection speed.

The YOLOv8 instance segmentation model (YOLOv8-Seg) is an improvement over the original YOLOv8 object detection framework. To enable pixel-level instance segmentation, this extension incorporates the YOLACT network [35]. For each object detected in the image, the model generates outputs that include a mask, class label, and confidence score. The modifications to the YOLOv8-Seg architecture were driven by the need to enhance its suitability for agricultural applications. The updated model incorporates advanced techniques to effectively address variations in object size and shape conditions. Enhanced segmentation capabilities ensure accurate identification of irregularly shaped

objects, which are essential for precision agriculture tasks. By refining feature extraction and fusion mechanisms, the improved YOLOv8-Seg achieves an optimal balance between speed and precision, making it highly effective for real-time agricultural monitoring. Figure 2 illustrates the network structure and details of YOLOv8-Seg.

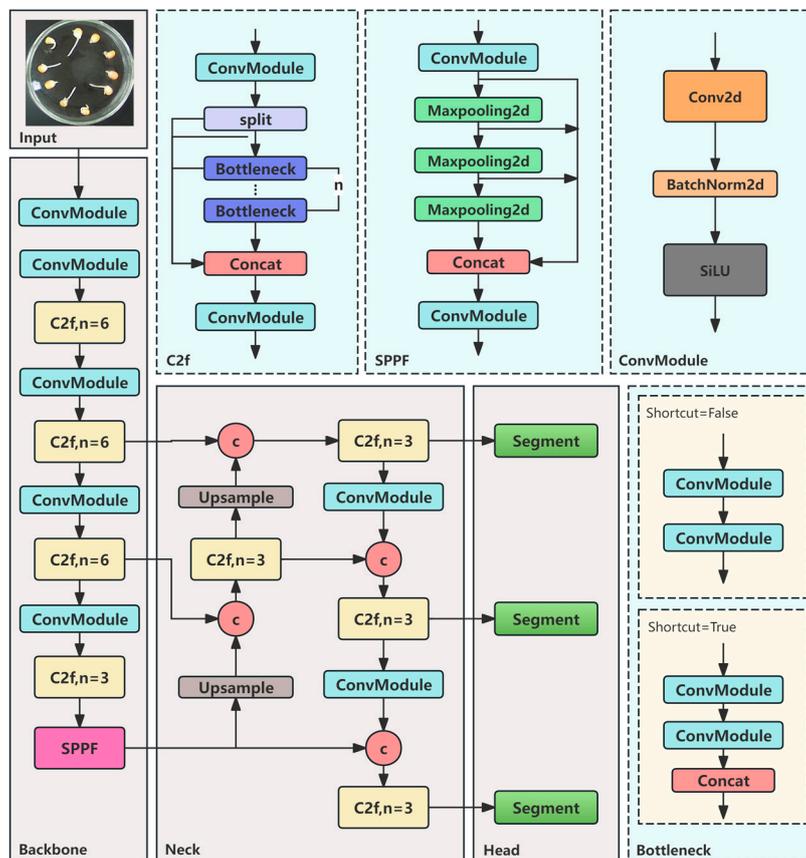
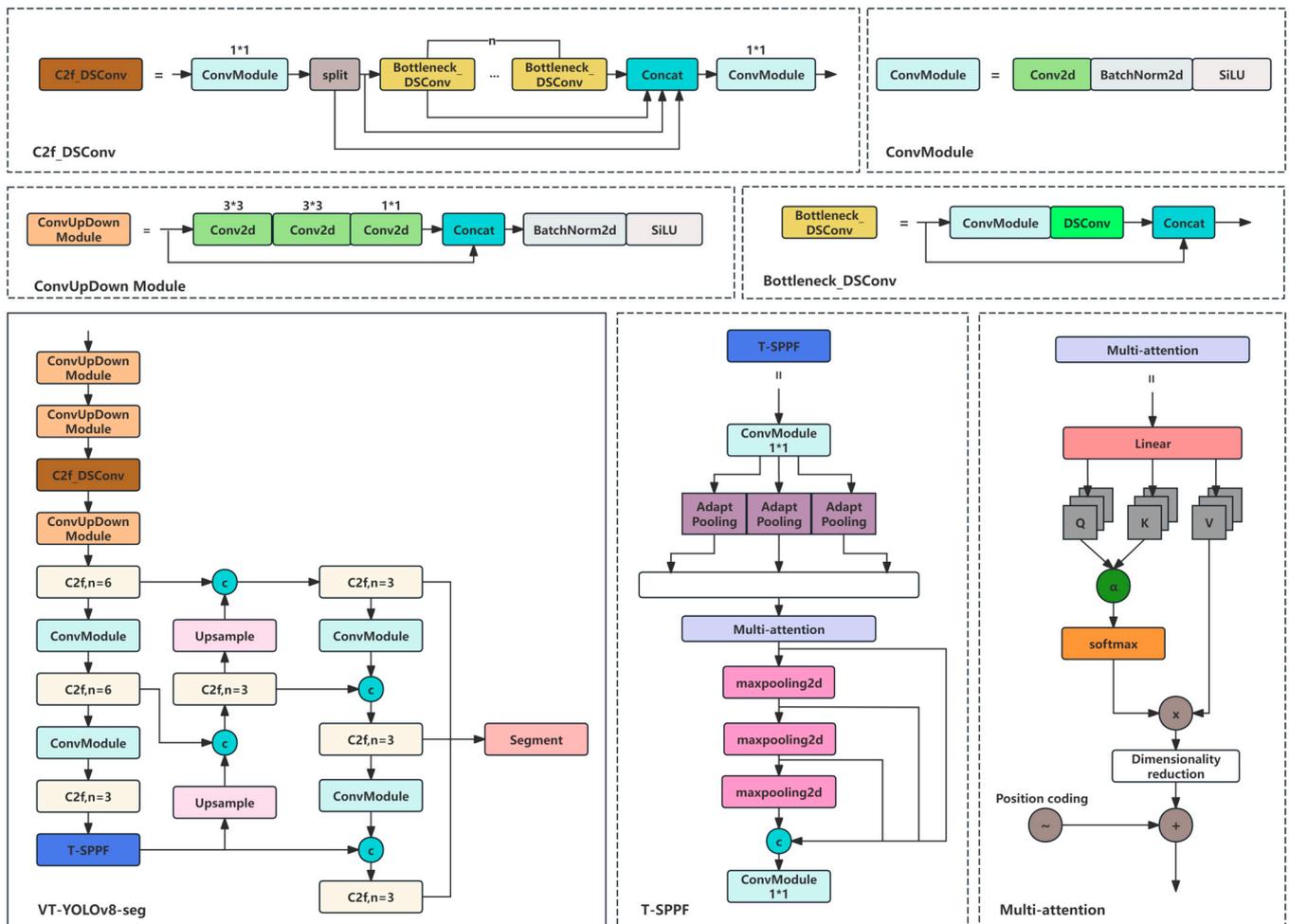


Figure 2. The model structure of the YOLOv8-Seg model.

### 2.3.2. Structure of the VT-YOLOv8-Seg Network

Seed evaluation is traditionally performed manually by researchers, a process that is labor-intensive, time-consuming, and susceptible to errors. This study proposes a segmentation method based on YOLOv8-Seg, named VT-YOLOv8-Seg, which not only detects germinating seeds but also separates the seed from the germ, enabling the computation of germ growth speed. This method minimizes error-prone manual steps and enhances the accuracy of seed bud segmentation, all while maintaining a compact and real-time model. Figure 3 presents the detailed network architecture.

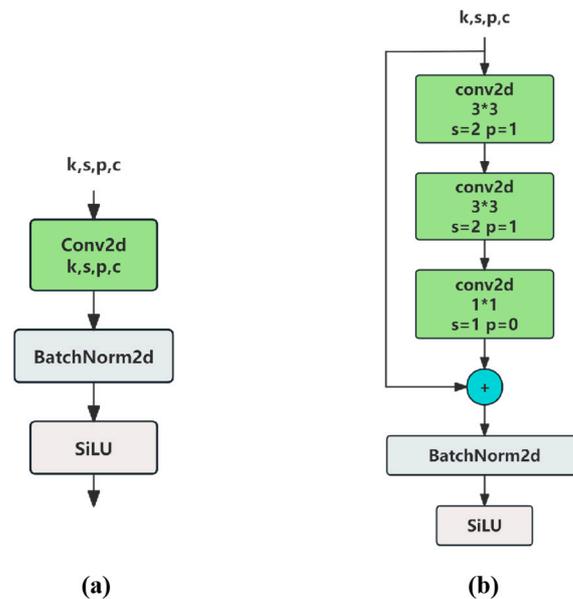


**Figure 3.** VT-YOLOv8-Seg network architecture diagram with all improvement points and some network details.

The VT-YOLOv8-Seg network structure consists of three primary components: the head, neck, and backbone. Before the image is input into the backbone, primary data preparation activities, such as data augmentation, are performed. The ConvUpDownModule refers to the modified convolution module, C2f\_DSConv denotes the C2f block incorporating DSConv, and T-SPPF is the improved SPPF module that introduces the transformer multi-head attention mechanism. These modules improve the base YOLOv8-Seg framework by enhancing feature extraction, segmentation precision, and computational efficiency. The ConvUpDownModule optimizes feature learning for datasets with homogeneous backgrounds, reducing redundancy and maintaining performance. DSConv enables accurate segmentation of irregular seed germs by refining edge detection, critical for distinguishing slender or tubular shapes. Finally, T-SPPF leverages self-attention mechanisms to capture multiscale features, enhancing both local and global contextual understanding. VT-YOLOv8-Seg enables accurate seed and germ segmentation in laboratory settings. The network parameters are nearly identical to those of the original model, with the addition of a network module. Its features ensure ease of use and real-time performance, and it can operate on standalone devices.

### 2.3.3. ConvUpDownModule

Figure 4a illustrates the convolutional block utilized in the YOLOv8 architecture. This study proposes the ConvUpDownModule, as shown in Figure 4b. The improved YOLOv8 architecture replaces the first three sets of  $3 \times 3$  convolutions in the backbone with the ConvUpDownModule. Leveraging the DenseNet and ResNet network architectures, two  $3 \times 3$  convolution kernels with  $s = 2$  and  $p = 1$  are first employed to reduce the size of the original feature maps to one-quarter of the input feature maps, followed by residual connections to facilitate feature learning.



**Figure 4.** (a) YOLOv8 convolutional block. (b) Structure of ConvUpDownModule. '3\*3' refers to a convolution kernel with a size of 3 by 3.

The maize seed germination dataset, collected in a laboratory setting, has a relatively homogeneous background. Reducing the feature map helps filter out redundant information while preserving essential features. A smaller feature map results in fewer parameters for subsequent layers to process, thereby accelerating both the training and inference processes.

### 2.3.4. C2f\_DSCConv

Accurately separating the seed from the germ is essential for determining the seed germination speed. The researchers' initial estimate of seed viability may contain mistakes due to inaccurate segmentation, which may have an impact on further breeding tests.

The square bounding box is indifferent to the local information of irregular targets, whereas the original YOLOv8-Seg network depends on the inner bounding box of the backbone's detection accuracy. Therefore, deformable convolution (DCNS) [36] was thought to be integrated into backbone to improve the network's segmentation effect on curly and slender maize germs. This would allow some of the  $3 \times 3$  convolution kernels to change their shapes to adapt to the irregular structure of the germs while keeping the convolution structure stable to reduce bias.

To create convolution kernels with a strong perception of irregular curves, dynamic snake convolution (DSCConv) [37] is introduced in this paper. It exhibits superior performance in tubular structure segmentation, adapts to elongated and twisted local structural features, and enhances the perception of geometrical structures. By employing strategies such as quantization and offsetting, DSCConv significantly reduces computational and

memory requirements, thereby accelerating the inference process. Figure 5 illustrates a variation in the convolution kernel.

This section describes the application of DSConv in extracting edge irregular local features, assuming 2D convolutional coordinates  $K$  and central coordinates  $K_i = (x_i, y_i)$ . The original  $3 \times 3$  convolution kernel  $K$  is denoted as

$$K = \{(x-1,y-1),(x-1,y),\dots,(x+1,y+1)\} \tag{1}$$

The convolution kernel is made more flexible by adding a deformation offset  $\Delta$ , which focuses on the irregular edges of tubular structures. The  $3 \times 3$  convolution kernel is expanded to a  $9 \times 9$  size and straightened along both axial directions using an iterative approach, as illustrated in Figure 6. The formula for each grid location in  $K$ , using the X-axis as an example, is  $K_{i\pm c} = (x_{i\pm c}, y_{i\pm c})$ , where  $c = \{0,1,2,3,4\}$  denotes the horizontal distance from the central grid. Each grid position  $K_{i\pm c}$  in  $K$  is determined through a cumulative process.  $K_{i\pm 1}$  increases by a bias  $\Delta = \{\delta|\delta \in [-1,1]\}$  relative to  $K_i$ , with the distance determined by the previous grid's position. Thus, the offset  $\Delta$  must be accumulated to ensure that the final convolution kernel  $K$  adheres to a linear structure.

The X-axis direction is

$$K_{i\pm c} = \begin{cases} (x_{i\pm c}, y_{i\pm c}) = \left( x_i + c, y_i + \sum_i^{i+c} \Delta y \right) \\ (x_{i\pm c}, y_{i\pm c}) = \left( x_i - c, y_i + \sum_{i-c}^i \Delta y \right) \end{cases} \tag{2}$$

The Y-axis direction is

$$K_{j\pm c} = \begin{cases} (x_{j\pm c}, y_{j\pm c}) = \left( x_i + \sum_j^{j+c} \Delta x, y_{j\pm c} \right) \\ (x_{j\pm c}, y_{j\pm c}) = \left( x_i + \sum_{j-c}^j \Delta x, y_{j\pm c} \right) \end{cases} \tag{3}$$

The offset  $\Delta$  is typically a small value, necessitating the conversion of position coordinates to integers through bilinear interpolation, as implemented below:

$$K = \sum_{K'} B(K', K) K' \tag{4}$$

where  $B$  represents the bilinear interpolation kernel,  $K$  denotes the fractional position in Equations (2) and (3), and  $K'$  represents the full spatial position.

As shown in Figure 6, DSConv spans a  $9 \times 9$  range during deformation, caused by shifts along two dimensions (X-axis and Y-axis). Leveraging its dynamic structure, DSConv adapts more effectively to elongated tubular structures, enhancing the perception of key properties.

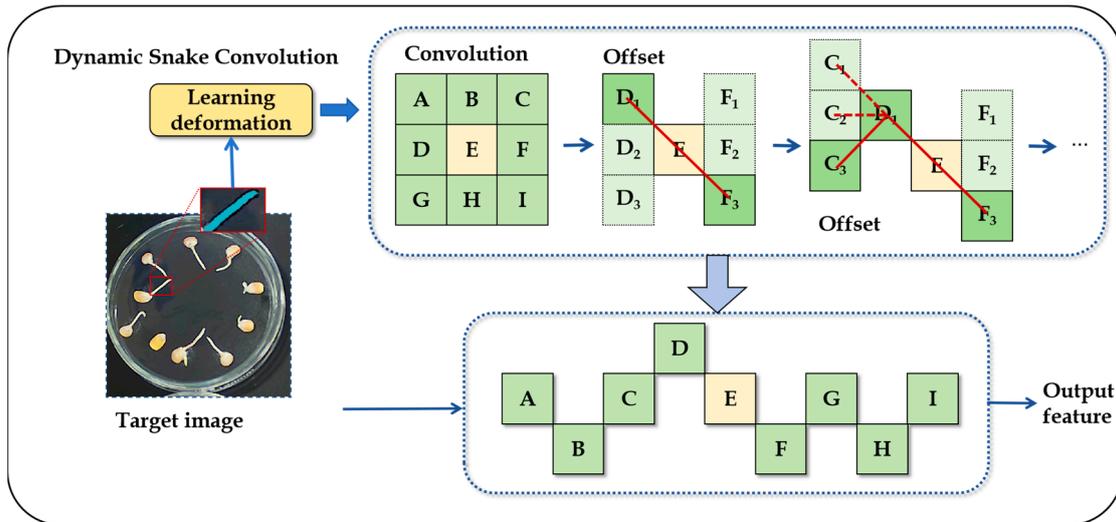


Figure 5. An example of DSConv convolution kernel variation [37].

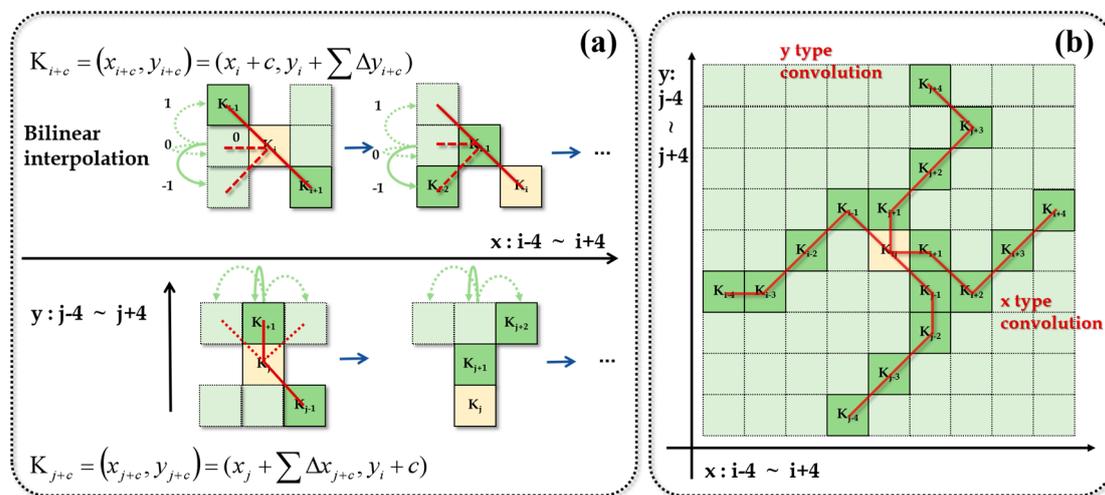
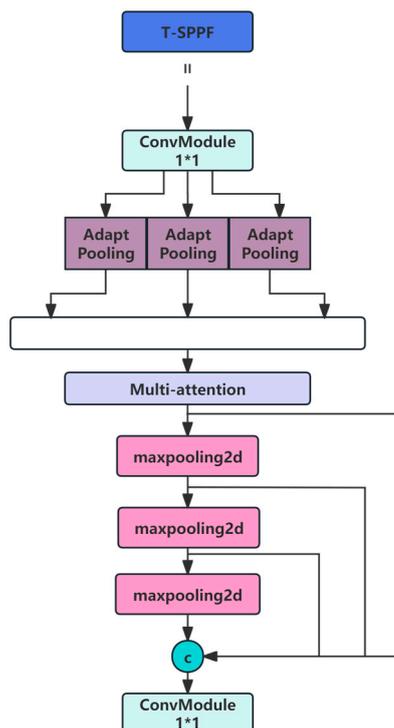


Figure 6. (a) Calculation method of DSConv coordinates. (b) Receptive field of DSConv [37].

### 2.3.5. T-SPPF

Compared to the seed, the newly germinated maize germ is small. The T-SPPF module is introduced to enhance segmentation accuracy across various scales, thereby capturing more and better feature information. Figure 7 illustrates the structure of the T-SPPF module.

The SPPF module first inputs the feature map through a  $1 \times 1$  convolutional block, which extracts the initial features while reducing computational complexity. To mitigate the impact of features with varying scale sizes, three parallel adaptive pooling layers are employed to adaptively adjust the pooling kernel size according to the size of the input features. After the splicing process, the fused features are passed into the multi-head attention mechanism module. By leveraging the correlation between pyramid feature maps, the multi-head attention mechanism enhances the ability to detect target segmentation across different scales, captures global information more effectively, and integrates the information learned by the head module at multiple scales [38]. Each attention head generates an attention weight matrix to weigh the input feature map. The outputs of all attention heads are subsequently combined into a single attention-weighted feature map. The transformer multi-head attention mechanism enhances the ability to capture global information, exploits correlations between pyramid feature maps, and strengthens the ability to detect segmentation targets at various scales.



**Figure 7.** T-SPPF module structure.

The input feature maps are then downsampled by three max-pooling layers, which reduce the feature map size and preserve critical information by selecting the maximum value in each pooling window. The T-SPPF module contains three max-pooling layers, each producing output feature maps with varying sizes and resolutions. The output feature maps from the three max-pooling layers undergo a splicing operation to combine them into a larger feature map. This spliced feature map contains feature information from multiple scales and locations, providing rich inputs for subsequent convolutional layers. Through multi-scale feature extraction and fusion, the model adapts more effectively to target objects of varying sizes, improving the accuracy of seed and germ segmentation.

#### 2.4. Experimental Parameters and Configuration

The model for this investigation was trained using the Windows 10 operating system. The neural network training model for the multi-class segmentation task was constructed using the PyTorch deep learning framework on a computer equipped with an Intel(R) Xeon(R) Gold 5218R CPU @ 2.10GHz and an NVIDIA GeForce RTX 2080 Ti GPU. Table 1 presents the most commonly used software version. The training setup includes a  $640 \times 640$  input image, a batch size of 16, an initial learning rate of 0.001, 50 epochs, and the Adam optimizer.

**Table 1.** Training and testing environment.

Configuration	Allocation
CUDA version	11.3
Python version	3.8
PyTorch version	1.12

### 2.5. Model Evaluation Criteria

Take the binary classification problem as an example in machine learning: if the actual result is positive and the predicted result is positive, the predicted result is positive (TP); if the actual result is negative, the predicted result is positive (FP); if the actual result is positive, the predicted result is negative (FN); and if the actual result is negative, the predicted result is negative (TN).

Precision (P) and recall (R) are derived from a confusion matrix that includes true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). Precision (P) is the proportion of predicted positive results to the total predicted positives, reflecting the network's classification accuracy, while recall represents the ratio of true positives correctly predicted by the network to the total number of actual true positives. The corresponding formulas are as follows:

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (5)$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (6)$$

Intersection over union (IoU) represents the ratio of the intersection to the union in object detection and is used to measure the extent of overlap between the model-generated boundaries and the original labeled boundaries. If the IoU exceeds 0.5, the object is deemed detectable. If the ground truth region is denoted as A and the labeled region as B, then

$$\text{IoU} = \frac{A \cap B}{A \cup B} \quad (7)$$

The average precision (AP) for a single category was calculated by ranking the model predictions based on their confidence scores and determining the area under the precision–recall (PR) curve as follows:

$$\text{AP} = \int_0^1 P(R) d(R) \quad (8)$$

Mean average precision (mAP) is the mean of the APs calculated across all categories, offering a comprehensive evaluation of the model's performance across all classes.

$$\text{mAP} = \frac{1}{n} \sum_{k=1}^{k=n} \text{AP}_k \quad (9)$$

Parameter (params) refers to the number of parameters that the model learns, which serves as a measure of its complexity. A larger number of parameters typically necessitates more GPU memory during training and influences the amount of storage and graphical memory the model occupies.

### 2.6. Indicators for Evaluating Seed Vigor

Seed vigor is a crucial predictor of a seed's germination potential, seedling emergence, and storage performance under various environmental conditions. The physiological maturation of seeds is influenced by both their genetic makeup and the conditions under which they are stored. The aging process, driven by the natural decline in seed vigor, affects seedling emergence and their ability to withstand adverse conditions. It also renders crops more susceptible to pests and diseases, ultimately influencing seed growth and yield potential.

Seed aging is a critical phase in the seed's life cycle. Physiological maturity represents the stage at which seeds are at their peak vitality, after which they begin to lose vigor. The

most commonly used method for simulating seed aging involves artificially aging treated seeds. The experiment utilized maize seeds of the same variety to investigate how seed aging influences maize growth and development. Maize seeds were artificially aged in an indoor incubator and greenhouse environment under conditions of high temperature and humidity. Three aging gradients were established: T1 (24 h), T2 (48 h), and T3 (72 h).

The aim of the experiment was to explore the relationship between seed aging and maize seed germination characteristics, as well as the changes in seed vigor under different aging conditions.

This study focused on four out of the twenty-seven germination metrics provided by the R package germinationmetrics [39]: mean germination time (MGT), germination index (GI), germination rate (seven days), and germination rate (three days). The algorithm used to calculate these metrics is presented in Table 2.

**Table 2.** Presentation of germination indicators.

	Formulas	Description
Final germination percentage (g)	$g = \frac{N_g}{N_t} * 100\%$	$N_g$ is the number of germinated seeds at moment $t$ after the start of the experiment, $N_t$ is the total number of seeds at moment $t$ after the start of the experiment.
Germination index (GI)	$GI = \sum \frac{N_g}{t}$	The sum of the number of germinated seeds at moment “ $t$ ” after the start of the experiment.
Mean germination time (MGT)	$MGT = \frac{\sum_{i=1}^k N_i T_i}{\sum_{i=1}^k N_i}$	where $T_i$ is the time from the beginning of the experiment to the “ $i$ ”th time interval, $N_i$ is the number of seeds germinated in the “ $i$ ”th time interval, and “ $k$ ”: is the total number of time intervals. ( $i = 24$ h, $k = 7$ )

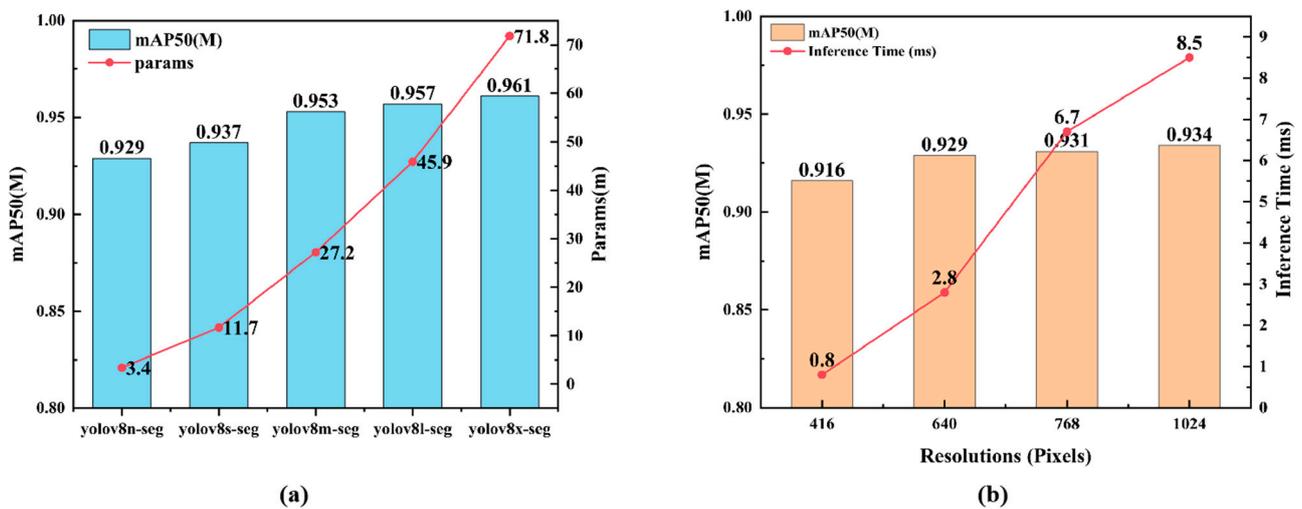
### 3. Results

#### 3.1. Merit-Seeking Experiment

YOLOv8n-Seg, YOLOv8s-Seg, YOLOv8m-Seg, YOLOv8l-Seg, and YOLOv8x-Seg represent the five YOLOv8 network models available on GitHub, each at different scales. To evaluate the segment mAP50 and the number of model parameters, we conducted experiments using YOLOv8-Seg models at different scales. The segment mAP50 values for the YOLOv8-Seg models at different scales are shown in Figure 8a. The results for the three models, YOLOv8m/l/x-seg, show no significant differences in terms of mAP50, the key metric for YOLO models, with values of 95.3%, 95.7%, and 96.1%, respectively. The YOLOv8n-Seg model has a size of 3.4M, and its segment mAP50 is 3.2% lower than that of the best-performing YOLOv8x-seg. YOLOv8n/s-Seg achieved segment mAP50 values of 92.9% and 93.7%, respectively. All results exceeded 90%. Considering the study’s use case, the need for deployment to edge devices, and the lightweight requirement, we selected YOLOv8n-Seg as the final experimental model.

During training, we used four different input image sizes:  $416 \times 416$  pixels,  $640 \times 640$  pixels,  $768 \times 768$  pixels, and  $1024 \times 1024$  pixels. This allowed us to assess the impact of image resolution on the segmentation performance of the YOLOv8n-Seg network. A

detailed summary of the effect of each image resolution on segment mAP50 and inference time is provided in Figure 8b. According to the results, the segment mAP50 increased by 1.3%, and the inference time rose by 2 ms when the image resolution was increased from  $416 \times 416$  pixels to  $640 \times 640$  pixels. This suggests that the slight increase in inference time is compensated by the improved instance segmentation performance of the model. However, the inference time increased by 5.1 ms and 7.7 ms, respectively, when the resolution was further increased to  $768 \times 768$  pixels and  $1024 \times 1024$  pixels. In comparison, the segment mAP50 showed only slight improvements of 0.15% and 0.18%, respectively. The findings indicate that the YOLOv8n-Seg network performs better when trained at a resolution of  $640 \times 640$  pixels.



**Figure 8.** (a) Segment mAP50 for YOLOv8-Seg models of different sizes. (b) Comprehensive overview of segment mAP50 and inference time for each image resolution.

### 3.2. Ablation Experiment

To verify the efficacy of the optimization techniques, ensure compatibility among different approaches, optimize resource utilization, and avoid confusion and inconsistencies during model training and decision-making, ablation comparison experiments were conducted in this study. Using the maize germination dataset, three distinct optimization techniques were incorporated into the original YOLOv8-Seg network. The YOLOv8-Seg (ConvUpDownModule), YOLOv8-Seg (C2f\_DSConv), YOLOv8-Seg (T-SPPF), and VT networks were compared, with each trained for 50 batches, as indicated in Table 3.

**Table 3.** YOLOv8-Seg performance combining various optimization techniques. Italicized text represents the improved modules of the model.

	P(Box)	R	mAP50	P(Mask)	R	mAP50	Params
YOLOv8n-Seg	0.922	0.891	0.931	0.914	0.904	0.926	11.80
YOLOv8n-Seg (ConvUpDownModule)	0.926	0.924	0.937	0.930	0.928	0.946	2.98
YOLOv8n-Seg (C2f_DSConv)	0.938	0.955	0.957	0.941	0.911	0.965	3.35
YOLOv8n-Seg (T_SPPF)	0.961	0.958	0.963	0.968	0.972	0.977	3.12
VT-YOLOv8n-Seg	0.973	0.979	0.985	0.972	0.977	0.982	3.15

The ConvUpDownModule was incorporated into the detection task, improving the model's accuracy, recall, and mAP50 to 92.6%, 92.4%, and 93.7%, respectively. These improvements are 0.4%, 3.3%, and 0.6% higher than those of the original model. While the model parameters were reduced by 0.22 million, the accuracy, recall, and mAP50 in the segmentation task increased to 93%, 92.8%, and 94.6%, respectively, which are 0.8%, 3.7%, and 2% higher than the original model. This demonstrates that the ConvUpDownModule can learn more useful information representations while using fewer parameters. It is also evident that the feature map's ability to extract features from maize seeds and germs in this study is not significantly affected by halving its size. Halving the feature map proves more effective than using the original feature map size.

The addition of DSConv increased segmentation accuracy by 2.7% and mAP50 by 3.9%, as well as detection accuracy by 1.6% and mAP50 by 2.6%, respectively, in comparison to the previous model. This suggests that dynamic snake convolution is effective at segmenting an object's edges, which is advantageous for maize germ segmentation. However, the improvement in recall is slight compared to the other metrics, and DSConv has limitations in recognizing different categories, which affects the model's balance; therefore, it should be combined with complementary methods.

Following the addition of the T-SPPF module, the model's accuracy, recall, and mAP50 in the detection task increased to 96.1%, 95.8%, and 96.3%, respectively, exceeding the original model by 3.9%, 6.7%, and 3.2%. In the segmentation task, the model's accuracy, recall, and mAP50 improved to 96.8%, 97.2%, and 97.7%, surpassing the original model by 5.4%, 6.8%, and 5.1%, respectively, while the number of parameters decreased by 0.08m.

The VT-YOLOv8-Seg model, integrating ConvUpDownModule, DSConv, T-SPPF, and YOLOv8-Seg, exhibited the best detection performance with an accuracy of 97.3%, a recall of 97.9%, and a mAP50 of 98.5%. In the segmentation task, the model achieved a mAP50 of 98.2%, a recall of 97.7%, and an accuracy of 97.2%. All metrics were significantly improved, while the number of parameters remained the same as in the original model.

### 3.3. Model Training Results

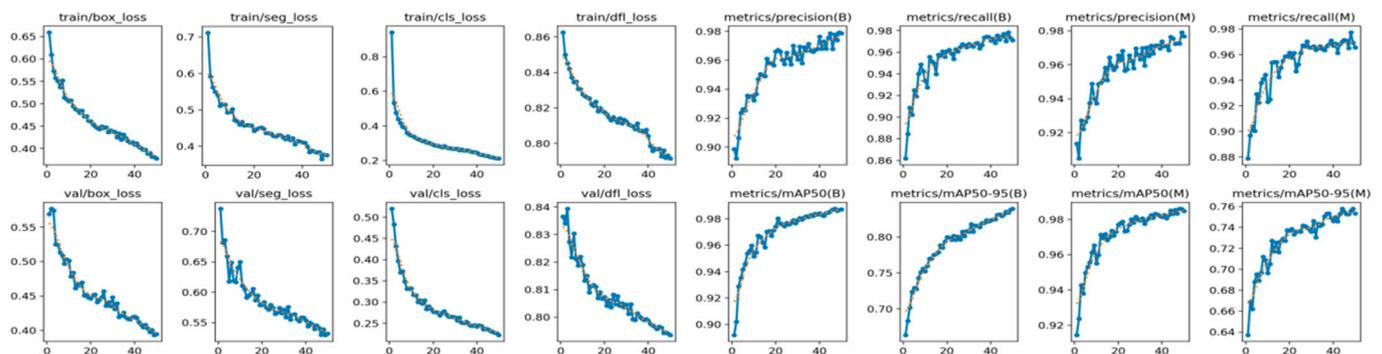
We trained VT-YOLOv8-Seg using the maize seed aging germination dataset, and Figure 9 illustrates the performance metrics during training and validation. During training, `box_loss` smoothly converges from an initial value above 0.4 to near 0, demonstrating high accuracy without overfitting. `seg_loss` also shows a clear decreasing and converging trend, while the reduction in classification loss (`cls_loss`) indicates an improvement in the model's reliability for target classification. The reduction in the direction/flow loss (`dfl_loss`) reflects the model's capacity to learn and predict the directional or flow-related characteristics of objects.

During model validation, all loss metrics show convergence trends similar to those observed during training, with final stable values slightly higher than the training results. This suggests that the model generalizes well to data outside the training set. With an average precision exceeding 0.9 at 50% IoU (mAP50) and a performance above 0.7 and more demanding 50–95% IoU (mAP50-95), mask precision (`precision(M)`) and recall (`recall(M)`) remain above 0.9. These results confirm the model's robust performance under stringent segmentation evaluation criteria.

The maize seed aging germination dataset was utilized to train VT-YOLOv8-Seg. Figure 9 illustrates the performance metrics for both training and validation. During the training process, `box_loss` gradually converges from an initial value above 0.4 to asymptotically close to 0, demonstrating high accuracy without overfitting. Additionally, `seg_loss` exhibits a clear downward and convergent trend, while the reduction in classification loss (`cls_loss`) indicates improved reliability in target classification. The model's

capacity to learn and predict object direction or flow is further reflected in the reduction in the specific loss metric, direction/flow loss (dfl\_loss).

During validation, all loss metrics exhibit convergence trends similar to those in training, with final stable values slightly higher than the training results. This suggests that the model generalizes well for unseen data outside the training set. With an average precision exceeding 0.9 at 50% IoU (mAP50) and a performance above 0.7 at the more challenging 50–95% IoU (mAP50-95), mask precision (precision(M)) and recall (recall(M)) remain above 0.9. These results confirm the model's robustness and its strong performance under stringent segmentation evaluation criteria.



**Figure 9.** Performance evaluation of YOLOv8-Seg model during training and validation. The blue solid line represents the value of a metric during the actual training/validation process, i.e., the real data points calculated for each epoch. The yellow dashed line is used to smooth the curve, making it easier to observe the overall trend.

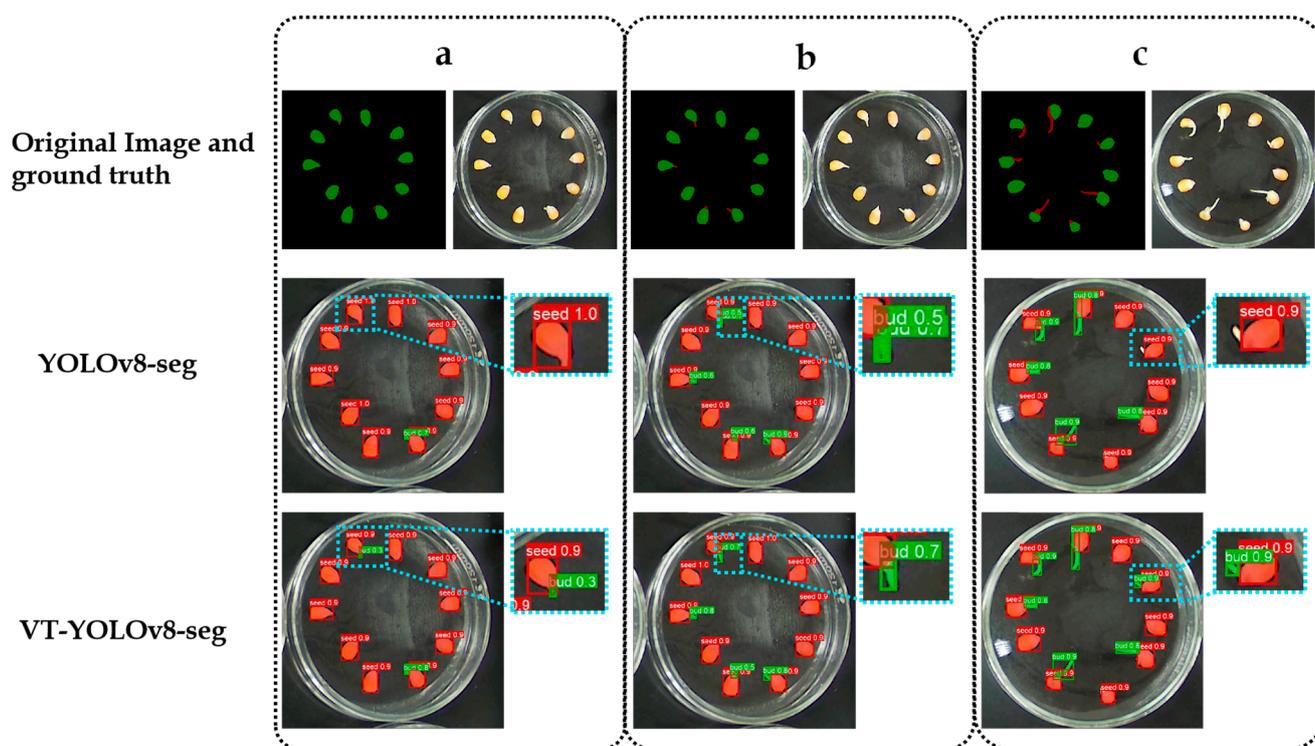
In conclusion, the VT-YOLOv8-Seg model demonstrates effective segmentation across various IoU thresholds, efficiently learning the features of maize seeds and shoots without overfitting and producing favorable training outcomes. Its low and stable loss values contribute to its strong performance on the maize seed aging germination dataset.

The original model (YOLOv8-Seg) exhibited limitations in detecting minor targets and failed to segment maize germs accurately, as shown by the mask comparison in Figure 10. The analysis reveals suboptimal detection and segmentation, particularly during the early growth stages of maize germs, limited capability in recognizing irregular edges, and insufficient convolutional flexibility. These shortcomings are primarily attributed to inadequate feature fusion, which hinders small target detection performance.

The enhanced model incorporates the ConvUpDownModule, which reduces model parameters and removes redundant information from feature maps; DSConv, featuring a flexible convolutional kernel that enhances the detection of irregular maize germ edges and improves edge segmentation accuracy; and the SPPF module, which manages multiscale feature information to enhance the network's ability to detect small targets. Tests demonstrate that our optimized model, VT-YOLOv8-Seg, effectively segments curved and irregular edges in seed and germ detection while handling small targets more proficiently compared to the original model.

When detecting the first stage of segmented maize germs, group (a) in Figure 10 compares the two models, showing that the improved VT-YOLOv8-Seg successfully detects small targets that the original model fails to identify. Group (a) illustrates how the improved VT-YOLOv8-Seg model successfully detects small germination targets that the original model fails to identify. Small targets are particularly challenging due to their low pixel density and indistinct boundaries in the early germination stage. The addition of the T-SPPF module enables better multiscale feature fusion, enhancing the model's sensitivity to these subtle features. In group (b), the original model's experimental results reveal

repeated detections, which are effectively resolved in the VT-YOLOv8-Seg model. This improvement is attributed to two key factors: the reduction in redundant background information and the enhanced capability to detect curved germ structures. By reducing the influence of irrelevant background features, the VT-YOLOv8-Seg model avoids misinterpreting overlapping or ambiguous areas as separate instances. Additionally, the strengthened ability to detect curved targets ensures more accurate identification and segmentation of bent germ structures, which are often misclassified by the original model. Missed detection remains an issue in group (c), with larger areas of missing buds compared to group (a). Missed detection remains a significant issue in group (c), where the original model shows larger areas of missing buds compared to the improved model in group (a). These omissions often occur for elongated and irregularly shaped targets, which the original architecture struggles to perceive due to limited edge adaptation capabilities. The integration of DSConv in the VT-YOLOv8-Seg model directly addresses this limitation. DSConv dynamically adjusts convolutional kernels to better align with the contours of elongated objects, significantly improving the segmentation accuracy for these challenging cases.



**Figure 10.** Groups (a)–(c) illustrate the model’s capability to detect and segment small targets in maize germs. Both group (a) and group (c) exhibit missed detection issues, with small germs in group (a) and elongated germs in group (c). Group (b), on the other hand, displays a problem of duplicate detections.

### 3.4. Comparative Experiment

We tested the VT-YOLOv8-Seg model with popular segmentation networks such as Mask-RCNN [40], YOLOv5-Seg [41], and YOLOv7-Seg [42] to confirm the effectiveness and benefits of the model. The VT-YOLOv8-Seg, the YOLOv8-Seg, the YOLOv7-Seg and the YOLOv5-Seg were evaluated on the test set after being trained with the same training parameters. For the Mask-RCNN, the learning rate, batch size, learning momentum, weight decay, and number of iterations were set to 0.004, 2, 0.9,  $1 \times 10^{-4}$ , and 30 epochs, respectively. As indicated in Table 4, each segmentation network’s precision (P), recall (R),

mean accuracy (mAP), and number of parameters were assessed during the entire experiment.

The enhanced YOLOv8-Seg algorithm's performance compared to other models is presented in Table 4. The accuracy, recall, and mAP50 are 97.3%, 97.9%, and 98.5% in the detection task and 97.2%, 97.7%, and 98.2% in the segmentation test, respectively. Compared to the YOLOv8-Seg method, the improvements are 5.1%, 8.8%, and 6.4% in the detection task and 5.8%, 7.3%, and 6.6% in the segmentation task. For the detection task, improvements are 7.8%, 10.3%, and 8.5% compared to the YOLOv5-Seg method, and for the segmentation task, improvements are 19.1%, 16.3%, and 12.8%, respectively. The detection task is improved by 9.6%, 12.9%, and 7% compared to the YOLOv7-Seg method, while the segmentation task is improved by 12.2%, 11%, and 7.6%. When compared to the Mask-RCNN method, the algorithm's evaluation metrics improved by 5.6%, 10.9%, and 6% for the detection task and by 4.6%, 8.3%, and 6.7% for the segmentation task.

In addition to its superior performance, VT-YOLOv8n-Seg maintained exceptional parameter efficiency, with only 3.15 million parameters. This is comparable to YOLOv8-Seg (3.20 million) but significantly lower than YOLOv7-Seg (81.5 million) and Mask-RCNN (239 million). The lightweight architecture of VT-YOLOv8n-Seg underscores its suitability for real-time applications on resource-constrained devices.

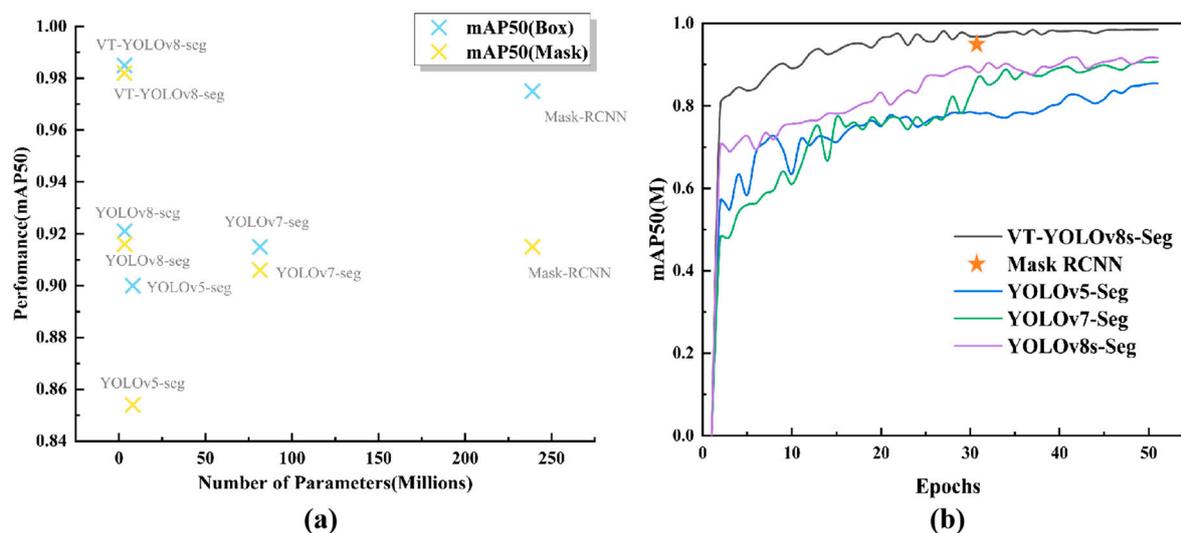
VT-YOLOv8n-Seg incorporates several innovative components to enhance performance. The ConvUpDownModule minimizes redundant feature maps, thereby improving computational efficiency and feature representation. The DSConv module is specifically designed to adapt to irregular germ edges, enabling accurate segmentation of curved and elongated structures. Leveraging a Transformer-based multi-head attention mechanism, the T-SPPF module introduces multi-scale feature fusion, enhancing the model's capacity to detect small and complex targets. Tailored for agricultural applications, VT-YOLOv8n-Seg strikes an optimal balance between high accuracy and low computational demands, making it ideal for deployment in resource-constrained environments. Its superior segmentation capabilities significantly enhance the tracking of seed germination, a critical factor in assessing seed vigor and improving agricultural productivity.

**Table 4.** Results for the five algorithms.

	<b>P(Box)</b>	<b>R</b>	<b>mAP50</b>	<b>P(Mask)</b>	<b>R</b>	<b>mAP50</b>	<b>Params</b>
Mask-RCNN	0.917	0.87	0.925	0.926	0.894	0.915	239
YOLOv5-Seg	0.895	0.876	0.90	0.781	0.814	0.854	7.9
YOLOv7-Seg	0.877	0.85	0.915	0.85	0.867	0.906	81.5
YOLOv8-Seg	0.922	0.891	0.921	0.914	0.904	0.916	3.20
VT-YOLOv8n-Seg	0.973	0.979	0.985	0.972	0.977	0.982	3.15

Figure 11a shows the relationship between the number of parameters and performance (mAP50) for Box and Mask detection across different models. The VT-YOLOv8-Seg model achieves the highest mAP50 scores for both Box and Mask detection while maintaining a significantly lower parameter count compared to Mask-RCNN. Specifically, VT-YOLOv8-Seg achieves an mAP50(B) and mAP50(M) close to 0.99, with approximately 3 million parameters, highlighting its efficiency and accuracy. In contrast, Mask-RCNN, with over 250 million parameters, delivers inferior performance in both metrics, reflecting a diminishing return in accuracy despite a substantially larger model size. The YOLOv5-Seg and YOLOv7-Seg models exhibit moderate performance, with mAP50(Box) and mAP50(Mask) values ranging between 0.9 and 0.92, while YOLOv8-Seg performs slightly better but still falls short of VT-YOLOv8-Seg. Figure 11b illustrates the training dynamics of mAP50(M) across epochs for the same models. VT-YOLOv8-Seg demonstrates superior convergence behavior, achieving a near-perfect mAP50(M) of 0.9 by the 50th epoch. This

rapid convergence reflects the effectiveness of the architectural enhancements, such as T-SPPF for multi-scale feature fusion and DSConv for improved detection of elongated and complex structures. In contrast, Mask-RCNN stabilizes at approximately 0.85 mAP50(M). YOLOv5-Seg and YOLOv7-Seg exhibit notable fluctuations during training and converge at approximately 0.8 and 0.9 mAP50(M), respectively. Although YOLOv8-Seg demonstrates improved stability compared to YOLOv5-Seg and YOLOv7-Seg, it remains inferior to VT-YOLOv8-Seg in both convergence speed and final performance.



**Figure 11.** Performance comparison of different models. (a) Relationship between the number of parameters (millions) and mAP50 for Box and Mask detection. (b) Training dynamics of mAP50(M) over 50 epochs.

### 3.5. Results of Germination Experiments on Aging Maize Seeds

#### 3.5.1. Detection Germinating Seeds

A comprehensive comparison was made between the variability of manual and platform counts to verify the accuracy of this seed germination model. The experiments were divided into four groups, recording the number of seeds sprouted every 24 h, to compare the modeled and manual germination counts. As shown in Figure 12, which compares manual counting and model-based detection results, the error in counting germinated seeds is relatively small, with a high degree of fit observed. The manual counting results were obtained by averaging the counts of five individuals to minimize subjectivity and enhance the reliability of the baseline measurements. Despite the small errors, the discrepancies are attributed to the current limitations in the model’s detection accuracy, which requires further refinement to enhance performance. These findings highlight the robustness of the VT-YOLOv8-Seg model while also suggesting areas for improvement.

Table 5, which compares the four seed groups, summarizes the findings of the four indicators. The results demonstrated that the mean germination time (MGT) exhibited an increasing trend, consistent with the hypothesis that seed vitality decreases as aging time increases. In contrast, the germination rate (G) and germination index (GI) indicators significantly decreased as the artificial aging treatment time was extended. Table 6 displays the indicator values for each of the four seed groups.

**Table 5.** Comparison of  $g(t=3)$ ,  $g(t=7)$ , GI, MGT values of four groups of seeds.

	$g(t=3)$	$g(t=7)$	GI	MGT
T0	36.6%	93.3%	16.08	5.586

T1	30.0%	86.7%	14.44	5.704
T2	0	83.3%	10.59	5.714
T3	0	76.7%	9.07	5.826

### 3.5.2. Germ Growth Speed

The germ area is determined by applying the best-trained parameters from VT-YOLOv8-Seg to the prediction task, counting the number of mask points, and plotting them on the graph while drawing the mask image. The total number of non-zero pixel points in the mask is counted, and the tensor is then converted into Python values to compute the mask's area post-prediction. The mask image distinguishes the seed and germ using two colors: green for the seed and red for the germ. Using the number of pixel points representing the germ's area, the Petri dish's pixel points, and the actual size of the Petri dish, the actual area of the germ can be calculated as follows:

$$\frac{\text{Actual area of germ}}{\text{Actual area of petri dish}} = \frac{\text{Total number of germ pixel points}}{\text{Total number of petri dish pixel points}} \quad (10)$$

A control group and three groups of maize seeds experiencing different lengths of aging were set up for germination experiments, with four groups, each consisting of 30 seeds (distributed across three Petri dishes), with seeds selected for consistency in position and morphology. A growth curve was fitted to the seeds using the mean germ area from the counted seeds. Most commercial automated seed vigor tests only assess germination within a three- to seven-day window. This study involved plotting bud growth curves of germ area versus time ratio by segmenting the seed buds and analyzing the results. It was observed that as the aging time gradient increased, the curve flattened and the germination speed slowed, as compared to the control and the three seed groups subjected to different aging gradients. An overall comparison of the growth curves for the four seed groups, shown in Figure 13, revealed that T3 (72 h aging time) had the slowest growth, T2 (48 h aging time) exhibited the latest time to first germination, and T0 (control) demonstrated the quickest growth and earliest germination. The curves obtained from the fitted exponential equation reflected the same trend as the original growth curve. After fitting the equation, the derived values were used to plot the derivative curve.

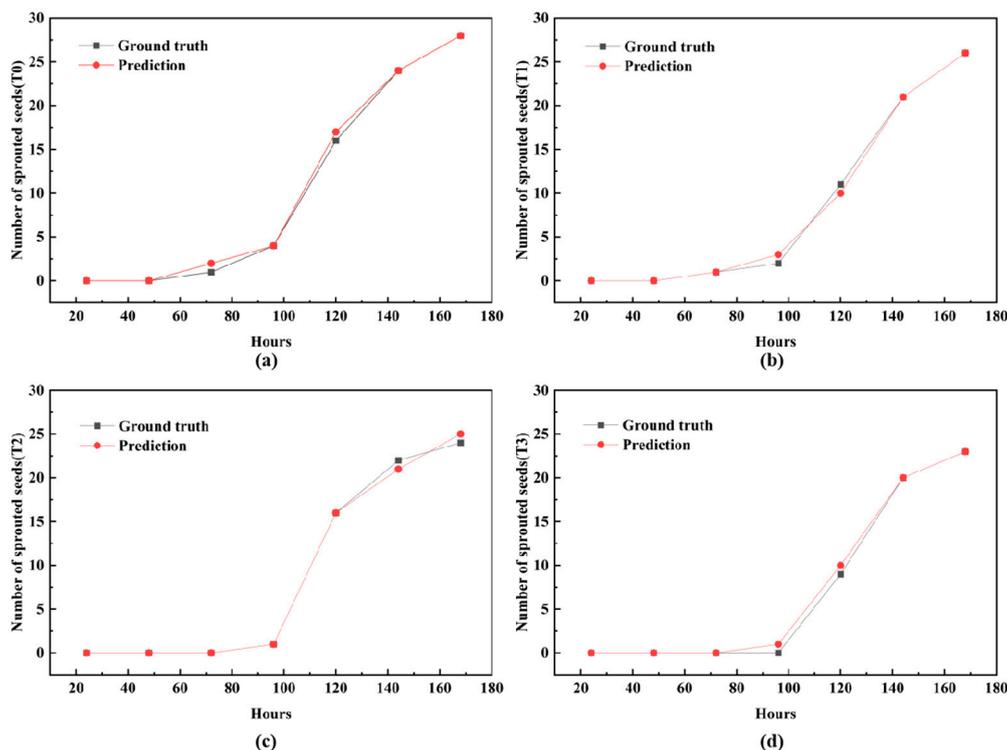


Figure 12. (a)–(d) compares the results between manual counting and model-based detection of maize seeds across four treatment groups with varying aging times.

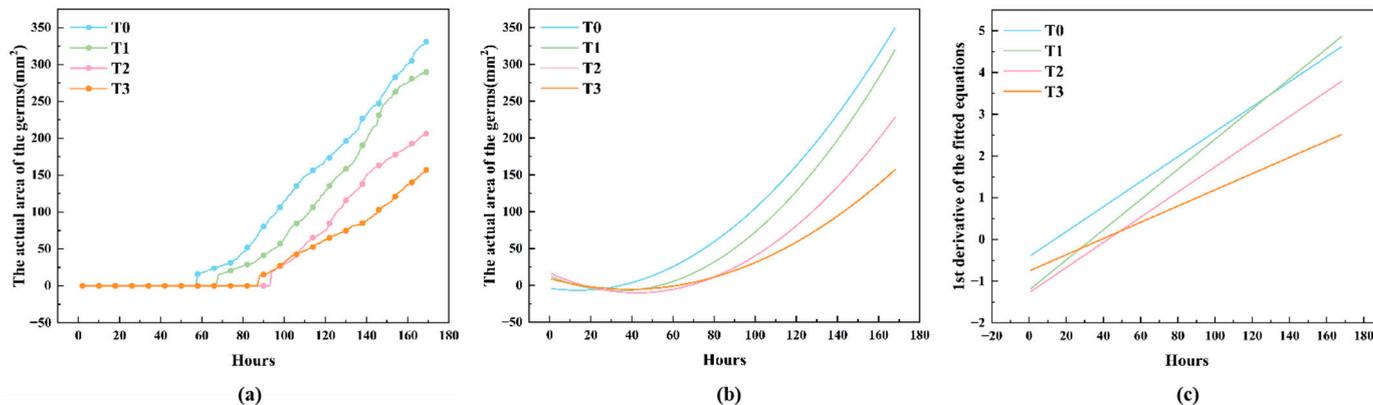


Figure 13. (a,b) The horizontal axis represents the germination experiment time, while the vertical axis indicates the maize germ areas, as calculated from the model outputs. (c) The horizontal axis represents the germination times, and the vertical axis displays the values of the first-order derivatives derived from curve fitting. (a) depicts the growth curve of the maize germ area over time. (b) compares the fitted curves of four sets. (c) presents the derivative equation obtained through first-order differentiation of the fitted equation, which approximates the growth rate of the maize germ. T0 (control), T1 (aging time 24 h), T2 (aging time 48 h), and T3 (aging time 72 h) are the experimental groups.

The growth curves of the four seed groups were analyzed by fitting equations using Origin software. After fitting the exponential equations, the data were processed using first-order differentiation to derive the derivative equations. The final derivative equations were derived to represent the germination rate. The germination velocity curve derived from the equation presents somewhat different results. The results show that T0 consistently demonstrated the highest germination velocity, outperforming other groups

throughout the experiment. T1, while slower initially, exceeded T0 in velocity after 130 h, suggesting that short-term high-temperature, high-humidity treatments may enhance seed vigor during the later stages of germination. T3 remained ungerminated before 80 h but exhibited a higher derivative rate than T2 during this period, which may be attributed to the limitations of the fitting process, indicating that the derivative model requires refinement for better precision. The analysis confirms that aging treatments reduce seed vigor, as evidenced by the flattened growth curves and lower germination rates in T2 and T3. However, the temporary recovery effect observed in T1 due to high humidity warrants further investigation. Minor discrepancies between the predicted and observed growth curves, particularly in T2 and T3, highlight the need for a more accurate germination rate equation tailored to this seed viability detection task. We plan to refine the model and propose a more accurate germination rate equation tailored to this seed viability detection task. The results of the fitted equations are presented in Table 6.

**Table 6.** The fitted equations with the corresponding values of  $R^2$ .

	Fitted Equations ( $y = ax^b$ )	$R^2$ (COD)
T0	$y_0 = 0.01352x^{2.3}$	0.9816
T1	$y_1 = 0.000661x^{2.8764}$	0.9877
T2	$y_2 = 0.000510x^{3.3133}$	0.9664
T3	$y_3 = 0.0006769x^{3.1866}$	0.9837

#### 4. Discussion

Seed vigor is traditionally assessed by examining germination rate, but calculating germination velocity provides a more nuanced approach to seed vigor analysis. In this study, 120 seeds were analyzed, with 30 seeds sampled per material. While this sample size balanced resource constraints and experimental feasibility, it limits generalizability. Specifically, the small sample size may not fully capture the variability within broader seed populations or diverse environmental conditions. Thus, the results should be interpreted with caution. Future studies should include larger and more diverse samples to enhance statistical robustness and broader applicability.

Advances in segmentation algorithms and model portability have made instance segmentation a valuable tool for evaluating both seed germination rate and velocity. For instance, YOLOv8-Seg has demonstrated significant potential in segmentation tasks. Reference [43] used YOLOv8-Seg as a base model, which was further improved for automatic segmentation of individual leaves in an image with a mAP50 of 86.4%. Reference [44] proposed improved YOLOv8 instance segmentation (BFFDC-YOLOv8-Seg) for accurately segmenting weeds in complex backgrounds. mAP50 was 98.8%, and these two examples show that the YOLOv8 family has excellent potential in segmentation tasks. These examples illustrate the efficacy of the YOLOv8 family in segmentation tasks.

The VT-YOLOv8-Seg model introduces an innovative approach for seed viability analysis, offering a quick, cost-effective, and efficient method for maize seed and germ segmentation. Despite significant advancements, further progress is necessary to improve the model's applicability across diverse crops. Key areas for improvement include the creation of a more comprehensive germ dataset to enhance feature diversity and generalization, as well as refinement of model accuracy, which is influenced by the pixel density in segmentation outputs. Moreover, to explore germination speed, future research should focus on developing algorithms capable of interpreting variations in germination dynamics, which could later be integrated into the model. Looking ahead, 3D image analysis could further advance our germ segmentation capabilities. Ultimately, these efforts aim to improve seed quality and food production, contributing to agricultural sustainability.

## 5. Conclusions

To enhance the accuracy of seed viability detection in maize, this study applies the YOLOv8-Seg network to the segmentation of maize germs. We propose the VT-YOLOv8-Seg model, which integrates dynamic snake convolution, ConvUpDownModule, and the T-SPPF mechanism to accurately delineate seed and germ boundaries, enabling precise observation of seeds' germination. VT-YOLOv8-Seg efficiently processes multi-scale feature information, improving detection of irregular germ edges, enhancing segmentation accuracy, and reducing model parameters by minimizing redundant feature map information.

The VT-YOLOv8-Seg model was used to count sprouted seeds, calculate the germination index, and assess the impact of aging gradients on maize seed germination vigor. A comparison of the germination counts from the VT-YOLOv8-Seg network with manual counting was performed to support the establishment of seed germination vigor classes.

By applying VT-YOLOv8-Seg to the segmentation of maize seeds and buds, the germination potential curve was derived, showing the ratio of germinal area to time. The slope of this curve indicates germination speed, providing a convenient method for monitoring seed vigor. This model offers a new analytical approach to measuring seed quality and serves as a technical reference for studying seed germination vitality and developing robust seed vitality rating models.

**Author Contributions:** Conceptualization, H.Y.; methodology, X.L.; software, X.L.; validation, Z.C.; formal analysis, C.B.; investigation, C.B.; resources, W.Z.; data curation, X.L.; writing—original draft preparation, X.L.; writing—review and editing, H.Y.; visualization, C.B.; supervision, W.Z.; project administration, H.Y.; funding acquisition, H.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** Please add: This study was supported by a major program of the Ministry of Agriculture and Rural Affairs (NK202302020205).

**Institutional Review Board Statement:** Not applicable.

**Data Availability Statement:** The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author(s).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Anbes, T.; Abebe, D.; Elmihun, M.; Asredie, S.; Ewenetu, W.; Chanie, Y. Assessment of seed quality and identifying sources of contaminants for maize seed across seed generations accessed through different seed source in Northwest Amhara. *Int. J. Food Agric. Nat. Resour.* **2024**, *5*, 85–96. <https://doi.org/10.46676/ij-fanres.v5i4.407>.
2. FAO. Database: Crops and Livestock Products FAOSTAT. Available online: <https://www.fao.org/faostat/en/#data/QCL> (accessed on 22 January 2025).
3. Filho, M.; Agricola, J.J.S. Seed vigor testing: An overview of the past, present and future perspective. *Sci. Agric.* **2015**, *72*, 363–374.
4. Finch-Savage, W.E.; Bassel, G.W. Seed vigour and crop establishment: Extending performance beyond adaptation. *J. Exp. Bot.* **2016**, *67*, 567–591. <https://doi.org/10.1093/jxb/erv490>.
5. King, T.; Cole, M.; Farber, J.M.; Eisenbrand, G.; Zabarar, D.; Fox, E.M.; Hill, J.P. Food safety for food security: Relationship between global megatrends and developments in food safety. *Trends Food Sci. Technol.* **2017**, *68*, 160–175. <https://doi.org/10.1016/j.tifs.2017.08.014>.
6. Milivojević, M.; Ripka, Z.; Petrović, T. ISTA rules changes in seed germination testing at the beginning of the 21st century. *J. Process. Energy* **2018**, *22*, 40–45.
7. Dell' Aquila, A. Digital imaging information technology applied to seed germination testing. A review. *Agron. Sustain. Dev.* **2009**, *29*, 213–221. <https://doi.org/10.1051/agro:2008039>.

8. Joosen, R.V.L.; Kodde, J.; Willems, L.A.J.; Ligterink, W.; van der Plas, L.H.W.; Hilhorst, H.W.M. GERMINATOR: A software package for high-throughput scoring and curve fitting of Arabidopsis seed germination. *Plant J.* **2010**, *62*, 148–159. <https://doi.org/10.1111/j.1365-313X.2009.04116.x>.
9. Hoffmaster, A.F.; Xu, L.; Fujimura, K.; Bennett, M.A.; McDonald, E.M.B. The Ohio State University Seed Vigor Imaging System (SVIS) for Soybean and Maize Seedlings. *Seed Technol.* **2005**, *27*, 7–24.
10. Liakos, K.G.; Busato, P.; Moshou, D.; Pearson, S.; Bochtis, D. Machine Learning in Agriculture: A Review. *Sensors* **2018**, *18*, 2674. <https://doi.org/10.3390/s18082674>.
11. Rehman, T.U.; Mahmud, M.S.; Chang, Y.K.; Jin, J.; Shin, J. Current and future applications of statistical machine learning algorithms for agricultural machine vision systems. *Comput. Electron. Agric.* **2019**, *156*, 585–605. <https://doi.org/10.1016/j.compag.2018.12.006>.
12. Castillo-Girones, S.; Munera, S.; Martínez-Sober, M.; Blasco, J.; Cubero, S.; Gómez-Sanchis, J. Artificial Neural Networks in Agriculture, the core of artificial intelligence: What, When, and Why. *Comput. Electron. Agric.* **2025**, *230*, 109938. <https://doi.org/10.1016/j.compag.2025.109938>.
13. Albayrak, U.; Golcuk, A.; Aktas, S.; Coruh, U.; Tasdemir, S.; Baykan, O. Classification and Analysis of *Agaricus bisporus* Diseases with Pre-Trained Deep Learning Models. *Agronomy* **2025**, *15*, 226. <https://doi.org/10.3390/agronomy15010226>.
14. Chaugule, M.A. Application of image processing in seed technology: A survey. *Technol. Adv. Eng.* **2014**, *2*, 153–159.
15. Dell’Aquila, A.; Van Eck, J.W.; van der Heijden, G.W.A. The application of image analysis in monitoring the imbibition process of white cabbage (*Brassica oleracea* L.) seeds. *Seed Sci. Res.* **2000**, *10*, 163–169.
16. Krubej, U.; Rozman, R.; Stajanko, D. The accuracy of the germination rate of seeds based on image processing and artificial neural networks. *Agricultura* **2015**, *12*, 19–14.
17. de Medeiros, A.D.; Pinheiro, D.T.; Xavier, W.A.; da Silva, L.J.; Dias, D. Quality classification of *Jatropha curcas* seeds using radiographic images and machine learning. *Ind. Crop. Prod.* **2020**, *146*, 7. <https://doi.org/10.1016/j.indcrop.2020.112162>.
18. Awty-Carroll, D.; Clifton-Brown, J.; Robson, P. Using *k*-NN to analyse images of diverse germination phenotypes and detect single seed germination in *Miscanthus sinensis*. *Plant Methods* **2018**, *14*, 7. <https://doi.org/10.1186/s13007-018-0272-0>.
19. Lurstwut, B.; Pornpanomchai, C. Image analysis based on color, shape and texture for rice seed (*Oryza sativa*, L.) germination evaluation. *Agric. Nat. Resour.* **2017**, *51*, 383–389.
20. Limao, M.A.R.; Dias, D.; Araújo, J.D.; Soares, J.M.; Nascimento, W.M.; da Silva, L.J. Classification of lentil seed vigor based on seedling image analysis techniques and interactive machine learning. *J. Seed Sci.* **2023**, *45*, 14. <https://doi.org/10.1590/2317-1545v45277692>.
21. Zhao, J.F.; Ma, Y.; Yong, K.C.; Zhu, M.; Wang, Y.Q.; Luo, Z.W.; Wei, X.; Huang, X.H. Deep-learning-based automatic evaluation of rice seed germination rate. *J. Sci. Food Agric.* **2023**, *103*, 1912–1924. <https://doi.org/10.1002/jsfa.12318>.
22. Huang, Z.L.; Wang, R.J.; Cao, Y.; Zheng, S.J.; Teng, Y.; Wang, F.M.; Wang, L.S.; Du, J.M. Deep learning based soybean seed classification. *Comput. Electron. Agric.* **2022**, *202*, 8. <https://doi.org/10.1016/j.compag.2022.107393>.
23. Eryigit, R.; Tugrul, B. Performance of Various Deep-Learning Networks in the Seed Classification Problem. *Symmetry* **2021**, *13*, 12. <https://doi.org/10.3390/sym13101892>.
24. Colmer, J.; O’Neill, C.M.; Wells, R.; Bostrom, A.; Reynolds, D.; Websdale, D.; Shiralagi, G.; Lu, W.; Lou, Q.J.; Le Maizeu, T.; et al. SeedGerm: A cost-effective phenotyping platform for automated seed imaging and machine-learning based phenotypic analysis of crop seed germination. *New Phytol.* **2020**, *228*, 778–793. <https://doi.org/10.1111/nph.16736>.
25. Nehoshtan, Y.; Carmon, E.; Yaniv, O.; Ayal, S.; Rotem, O. Robust seed germination prediction using deep learning and RGB image data. *Sci Rep* **2021**, *11*, 10. <https://doi.org/10.1038/s41598-021-01712-6>.
26. Reddy, D.R.; Cheruku, R.; Kodali, P. SeedAI: A novel seed germination predictionsystem using dual stage deep learning framework. *Environ. Res. Commun.* **2023**, *5*, 14. <https://doi.org/10.1088/2515-7620/ad16f1>.
27. Genze, N.; Bharti, R.; Grieb, M.; Schultheiss, S.J.; Grimm, D.G. Accurate machine learning-based germination detection, prediction and quality assessment of three grain crops. *Plant Methods* **2020**, *16*, 11. <https://doi.org/10.1186/s13007-020-00699-x>.
28. Fu, X.Q.; Han, B.; Liu, S.Y.; Zhou, J.Y.; Zhang, H.W.; Wang, H.B.; Zhang, H.; Ouyang, Z.Q. WSVAS: A YOLOv4-based phenotyping platform for automatically detecting the salt tolerance of wheat based on seed germination vigour. *Front. Plant Sci.* **2022**, *13*, 13. <https://doi.org/10.3389/fpls.2022.1074360>.
29. Yao, Q.; Zheng, X.M.; Zhou, G.M.; Zhang, J.H. SGR-YOLO: A method for detecting seed germination rate in wild rice. *Front. Plant Sci.* **2024**, *14*, 16. <https://doi.org/10.3389/fpls.2023.1305081>.
30. Sun, W.; Xu, M.; Xu, K.; Chen, D.; Wang, J.; Yang, R.; Chen, Q.; Yang, S. CSGD-YOLO: A Maize Seed Germination Status Detection Model Based on YOLOv8n. *Agronomy* **2025**, *15*. <https://doi.org/10.3390/agronomy15010128>.

31. Terven, J.; Cordova-Esparza, D.M.; Romero-Gonzalez, J.A. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. *Mach. Learn. Knowl. Extr.* **2023**, *5*, 1680–1716.
32. Chen, C.; Bai, M.; Wang, T.; Zhang, W.; Yu, H.; Pang, T.; Wu, J.; Li, Z.; Wang, X. An RGB image dataset for seed germination prediction and vigor detection - maize. *Front. Plant Sci.* **2024**, *15*. <https://doi.org/10.3389/fpls.2024.1341335>.
33. Liu, S.; Qi, L.; Qin, H.; Shi, J.; Jia, J. Path Aggregation Network for Instance Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018.
34. Lin, T.Y.; Dollar, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature Pyramid Networks for Object Detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
35. Bolya, D.; Zhou, C.; Xiao, F.; Lee, Y.J. YOLACT: Real-Time Instance Segmentation. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 9156–9165.
36. Dai, J.; Qi, H.; Xiong, Y.; Li, Y.; Zhang, G.; Hu, H.; Wei, Y. Deformable Convolutional Networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017.
37. Qi, Y.; He, Y.; Qi, X.; Zhang, Y.; Yang, G. Dynamic Snake Convolution based on Topological Geometric Constraints for Tubular Structure Segmentation. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Paris, France, 4–6 October 2023.
38. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. In Proceedings of the 31st Conference on Neural Information Processing Systems Long Beach, CA, USA, 4–9 December 2017.
39. Aravind, J.; Devi, S.V.; Radhamani, J.; Jacob, S.R.; Srinivasan, K.; ICARNBGPR. *Germination Metrics: Seed Germination Indices and Curve Fitting*; Available online: <https://github.com/cran/germinationmetrics> (accessed on 22 January 2025).
40. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. In Proceedings of the IEEE International Conference on Computer Vision; Venice, Italy, 22–29 October 2017; pp. 2961–2969.
41. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the Computer Vision & Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016.
42. Wang, C.Y.; Bochkovskiy, A.; Liao, H.Y.M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022.
43. Wang, P.; Deng, H.; Guo, J.X.; Ji, S.Q.; Meng, D.; Bao, J.; Zuo, P. Leaf Segmentation Using Modified YOLOv8-Seg Models. *Life* **2024**, *14*, 14. <https://doi.org/10.3390/life14060780>.
44. Lyu, Z.; Lu, A.J.; Ma, Y.L. Improved YOLOv8-Seg Based on Multiscale Feature Fusion and Deformable Convolution for Weed Precision Segmentation. *Appl. Sci.* **2024**, *14*, 17. <https://doi.org/10.3390/app14125002>.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.