



Segmentation of Morphological Structures of Soybean Seedlings Using Convolutional Neural Networks

Pedro Henrique dos S. e Silva¹, Emannuel Diego G. de Freitas^{1,2}, José Luciano M. Neto², Haynna F. Abud³, Danielo G. Gomes¹

¹Departamento de Engenharia de Teleinformática, Universidade Federal do Ceará, Fortaleza-CE, Brasil.

²Instituto Federal de Educação Ciência e Tecnologia do Ceará, campus Iguatu, Iguatu-CE, Brasil.

³Image Pesquisas, Av. Humberto Monte, Bloco 310, Galpão 17, Fortaleza-CE, Brasil.

**A paper from the Proceedings of the
17th International Conference on Precision Agriculture and 11th
Brazilian Congress on Precision and Digital Agriculture
13-15 July 2026
Porto Alegre, Brazil**

Abstract.

Soybean is one of the most relevant and highly demanded agricultural commodities worldwide, playing a central role in the production chains of food, animal feed, and vegetable oil. Globally, there is a growing demand for food, which has driven strategies to meet the needs of the world population. Many of these strategies involve expanding arable land and excessively exploiting soil nutrients in an unsustainable manner. An alternative strategy to address this issue is to incorporate seed lot quality assessment techniques into production systems. In this way, producers are provided with a quantifiable estimate of the individual germination capacity of each lot and the quality of their inputs, promoting a significant increase in production without the need to expand cultivated areas. This quality control process can be conducted through germination and vigor tests. Vigor tests are more sensitive in detecting differences in the physiological performance of seeds, providing complementary information that is often more indicative of field emergence potential than standard germination tests. Manual evaluation of a seed lot requires considerable time from the analyst; however, it is common for dozens, hundreds, or even thousands of samples to be analyzed, a volume that generates excessive workload, low efficiency, and various errors in routine operations. To mitigate these issues, several computer vision techniques were explored in this study to assist in vigor analysis. The dataset was built and developed using the Roboflow platform, and data annotation was limited to three classes: cotyledon, hypocotyl, and primary root, which represent the fundamental morphological structures of soybean seedlings. A total of 80 RGB images with a resolution of 4000 x 3000 pixels were acquired using a smartphone camera positioned on a fixed support, while the specimens were arranged on specific paper used for the germination test, with bottom lighting. The dataset was divided into 70% for training, 20% for validation, and 10% for testing, with augmentation applied only to the training partition. The development environment consisted of an Intel® Xeon® CPU @ 2.20 GHz processor, 52 GB of RAM, and an NVIDIA T4 GPU with 15 GB of VRAM. The algorithms used to identify and separate these classes in the images rely on artificial intelligence. Among the evaluated methods, YOLOv9, configured with a learning rate of 0.001, batch size of 8, the AdamW optimizer, and trained for 100 epochs, demonstrated the best performance,

The authors are solely responsible for the content of this paper, which is not a refereed publication. Citation of this work should state that it is from the Proceedings of the 17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture. EXAMPLE: Last Name, A. B. & Coauthor, C. D. (2026). Title of paper. In Proceedings of the 17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture (unpaginated, online). Monticello, IL: International Society of Precision Agriculture and Brazilian Congress on Precision and Digital Agriculture.

achieving precision values of 98%, 78%, and 88%, and recall values of 97%, 72%, and 75% for the cotyledon, hypocotyl, and primary root classes, respectively. These variations in the metrics are due to the complexity of each structure and its representativeness in the image. This study proposes alternatives capable of identifying the main morphological structures of soybean seedlings and reinforces the potential of artificial intelligence as a tool to automate and enhance agricultural processes that have traditionally depended on manual evaluation.

Keywords.

soybean seedling, computer vision, image segmentation, deep learning, machine learning.

Introduction

Human population growth has been driving a steep global demand for food (MAJA; AYANO, 2021; KOPITKE et al., 2019). In response to this demand, global agricultural production has reached significant levels in recent decades. According to the Food and Agriculture Organization of the United Nations (FAO) (2024), global primary crop production totaled approximately 9.9 billion tons, representing an increase of 27% compared to 2010. According to the Brazilian Agricultural Research Corporation (EMBRAPA) (2021), Brazil has consolidated its position in recent decades as one of the leading actors in global agribusiness.

According to EMBRAPA (2021), soybean plays a central role in this performance. In 2020, Brazil became the world's largest producer, with 126 million tons, representing more than one-third of global production. Historically, Brazil's share of global soybean production reached 28.2%, a position sustained over decades and surpassed only by the United States. In international trade, Brazilian dominance is also evident: in 2020, the country was the world's largest soybean exporter, accounting for approximately 50% of the global market. Over the last 21 years, Brazil has been responsible for more than 39% of global exports, generating a value of 30 billion US dollars in 2020 alone and 346 billion US dollars during the first two decades of the century.

However, although arable land is expected to increase by approximately 7% by 2030, there is limited availability of land for agricultural expansion without large-scale deforestation or the use of marginal lands (SMITH et al., 2010). These practices may intensify environmental degradation and create additional challenges for soil health, climate security, and water resources (MAJA; AYANO, 2021; KOPITKE et al., 2019; SILVA et al., 2024). In this scenario, meeting the growing demand for food sustainably requires strategies that go beyond simply expanding cultivated areas. Therefore, optimizing resource allocation and increasing productivity in already established agricultural areas become essential.

In this context, one of the most effective strategies for increasing agricultural productivity is the use of high-vigor seeds (MEDEIROS et al., 2020; ZHANG et al., 2023; ZOU et al., 2023). From this perspective, producers require seeds with high physiological potential, capable of ensuring adequate crop establishment and promoting rapid and uniform initial plant development, as these attributes are decisive for achieving high productivity levels and commercial success in crop production (FILHO, 2005; HÖFS et al., 2004; NAKAGAWA et al., 1985).

The physiological potential of seeds is traditionally evaluated through germination and vigor tests, which constitute essential tools for quality control in seed production (Ministério da Agricultura, Pecuária e Abastecimento, 2009a). These tests make it possible to identify seed lots with greater field establishment potential, contributing to higher efficiency and reliability in the production process (TEKRONY; EGLI, 1991; VANZOLINI; CARVALHO, 2002).

Despite their importance, many of these tests are based on visual evaluation performed by trained analysts. Although well established in laboratory routines, this process may present limitations inherent to human interpretation, such as subjectivity, variability among evaluators, and fatigue throughout the analyses due to the considerable time required to examine large numbers of samples. These factors may compromise the standardization and reproducibility of results

(DELL'AQUILA, 2009; MARCOS-FILHO, 2015).

Given these limitations, opportunities arise for the use of automated methods capable of reproducing visual tasks with greater precision and consistency. In this context, Computer Vision stands out as a promising technology, as it enables the extraction, processing, and interpretation of information contained in images in a rapid, objective, and scalable manner. Its application to seed analysis can transform traditionally manual procedures into automated processes, contributing to greater efficiency and reliability in vigor assessment (CASTAN et al., 2018; CONRAD, 1997; HOFFMASTER et al., 2003b; MCNERTNEY, 1999a).

However, for Computer Vision systems to extract reliable information and perform quantitative analyses from images, the proper segmentation of regions of interest is essential (GONZALEZ; WOODS, 2018). This process consists of the precise division of a digital image into its different constituent parts, separating objects from the background and from one another based on neighborhood similarity or the identification of discontinuities (ZHANG; LI, 2014; ANAND et al., 2022). Segmentation is recognized as one of the most critical stages in Computer Vision systems, since the quality of this stage directly influences the overall performance of the system (XU, 2024). Therefore, accurate and consistent segmentation is an indispensable requirement for proper feature extraction and object classification.

To contribute to this demand, the present study aims to evaluate Deep Learning algorithms, including Mask Region-based Convolutional Neural Network (Mask R-CNN), U-Net, and architectures from the You Only Look Once (YOLO) family. The comparative analysis seeks to identify which approaches achieve superior performance in the semantic segmentation of soybean seedling parts (cotyledon, hypocotyl, and radicle), providing support for the selection of more robust and accurate models for applications in seed technology.

Deep Learning Architectures

Mask R-CNN, proposed by He et al. (2018), is one of the most influential and widely used architectures for instance segmentation tasks, combining classification, object detection, and pixel-level segmentation within a single model. This architecture is based on a two-stage pipeline: (i) the generation of Regions of Interest (RoIs) through a Region Proposal Network (RPN) and (ii) the refinement of these regions for class prediction and bounding box regression, as shown in Figure 1. The core innovation of Mask R-CNN lies in the addition of a third parallel branch responsible for predicting the binary masks corresponding to each detected instance.

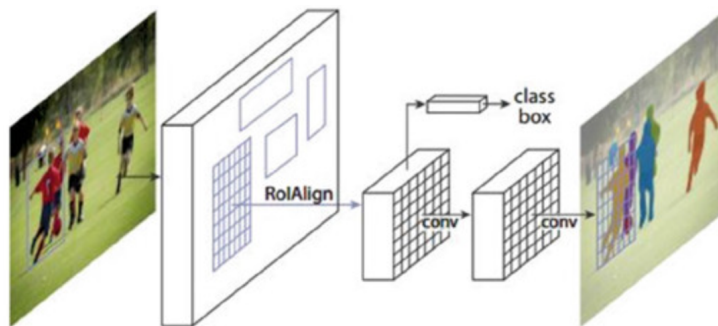


Fig 1. Mask R-CNN architecture for instance segmentation proposed by He et al. (2018, p. 1).

U-Net, proposed by Ronneberger, Fischer, and Brox (2015), is one of the most influential architectures for semantic segmentation tasks, particularly in domains requiring high spatial precision, such as biomedical imaging. Its main structure is composed of two symmetrical paths: a contraction path (encoder) (KADRY et al., 2025), responsible for feature extraction and the progressive reduction of spatial resolution, and an expansion path (decoder) (BRADSHAW; MCMILLAN, 2021), which reconstructs the spatial resolution, giving rise to the characteristic “U” shape of this architecture, as illustrated in Figure 2. The distinguishing feature of U-Net is that, in

the expansion path, each upsampling stage is combined with high-resolution feature maps from the contraction path through mechanisms known as skip connections (OUGUZ; ERTUUGRUL, 2023), reducing information loss throughout the encoder–decoder pathway.

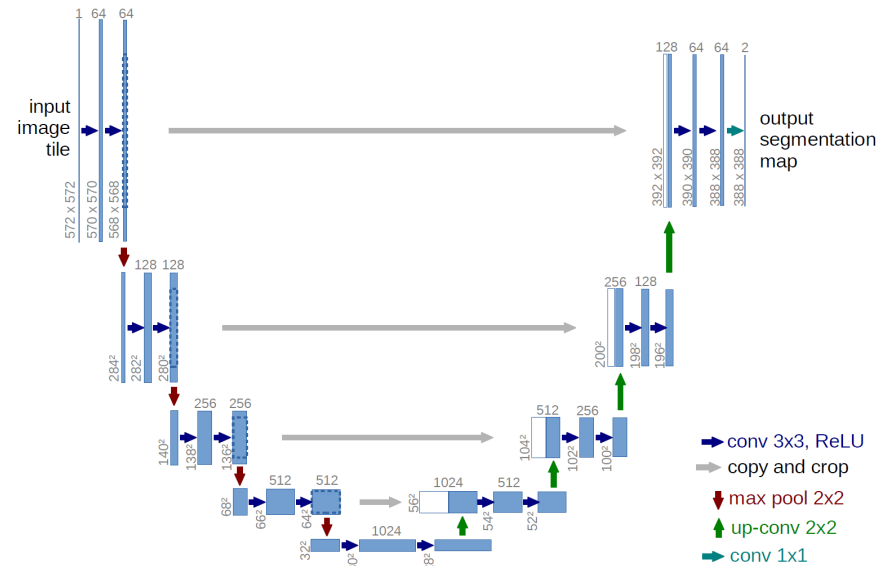


Fig 2. Classical U-Net architecture proposed by Ronneberger, Fischer, and Brox (2015, p. 2).

The YOLO architecture emerged as a response to the speed limitations of two-stage object detectors. While methods such as R-CNN (GIRSHICK et al., 2014) and Faster R-CNN (GIRSHICK, 2015)—architectures that preceded Mask R-CNN—performed a separate region proposal stage, YOLO introduced the concept of processing the entire image in a single pass, directly predicting bounding boxes, classes, and confidence scores using a single convolutional network, while being sufficiently fast for real-time applications (Redmon et al., 2016).

YOLO-v8, introduced by Jocher, Chaurasia, and Qiu (2023), is based on CSPDarknet-53, which employs Cross Stage Partial (CSP) connections to improve gradient flow and computational efficiency. Its neck combines the Path Aggregation Network (PANet) with Spatial Pyramid Pooling – Fast (SPPF) to aggregate features at multiple scales, while the head preserves a dense structure with separate branches for classification and regression. This architecture is illustrated in Figure 3.

The most important change is the adoption of an anchor-free paradigm (already present in some of its predecessors), in which the model directly predicts object centers without relying on predefined anchor boxes. This modification simplifies training, improves generalization, and reduces the number of generated bounding boxes, thereby decreasing the computational cost of Non-Maximum Suppression (NMS) and accelerating inference.

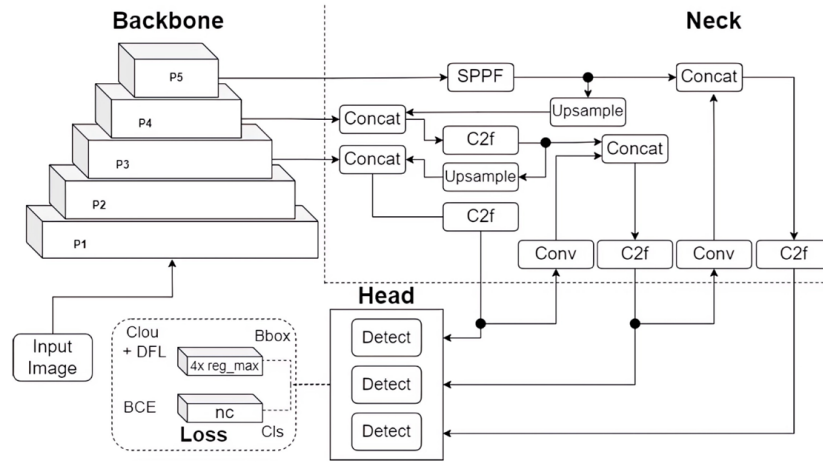


Fig 3. YOLO-v8 architecture proposed by Bento, Paixão, and Alvarez (2025, p. 5).

The YOLO-v9 architecture (Wang and Liao, 2024) was developed to overcome limitations related to information loss and gradient instability observed in previous architectures. The main advancement introduced is the Programmable Gradient Information (PGI) module, designed to improve gradient reliability during training without increasing inference cost.

PGI combines a main branch (used during inference), an auxiliary reversible branch, and a multi-level information pathway, ensuring improved gradient flow and more stable learning. These structures are presented in Figure 4 as the main branch, auxiliary reversible branch, and multi-level auxiliary information.

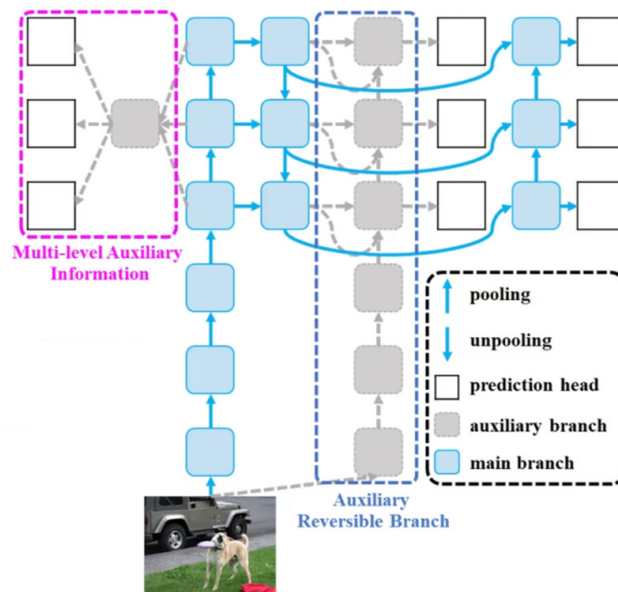


Fig 4. PGI of the YOLOv9 architecture adapted from Wang, Yeh, and Liao (2024, p. 5).

Complementing PGI, YOLO-v9 introduces the Generalized Efficient Layer Aggregation Network (GELAN), a modular architecture that enhances feature aggregation and optimizes gradient propagation. Based on the principles of the Extended Efficient Layer Aggregation Network (ELAN) and Cross Stage Partial Network (CSPNet), GELAN employs standard convolutions while maintaining an efficient, flexible design suitable for different model scales. These modifications and architectural enhancements can be observed in Figure 5.

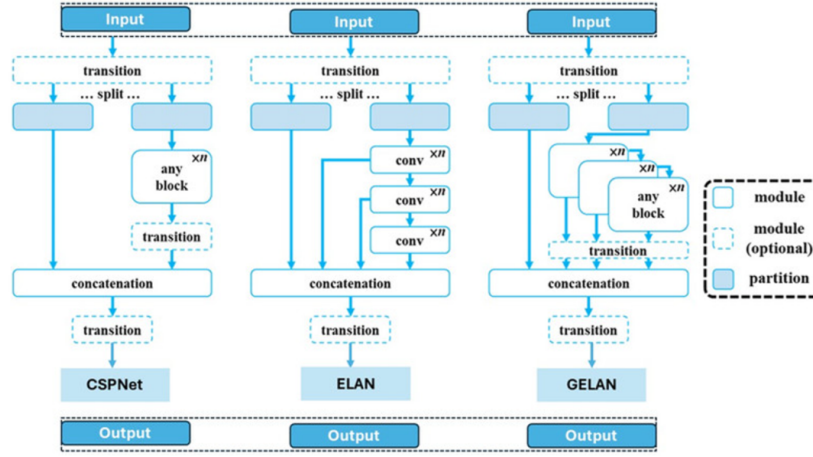


Fig 5. YOLOv9 architecture incorporating CSPNet, ELAN, and GELAN modules from Jegham et al. (2024, p. 6).

Materials and Methods

Dataset

The image dataset selected for the development of the proposed study consisted of a total of 80 images of soybean seedling cultivars. The original images have a resolution of 4000 x 3000 pixels and include approximately 50 seedlings with three days of development, as well as some non-germinated seeds, reflecting the typical variability observed in commercial seed lots. Figure 6 presents some sample images from the dataset.

The seedling image annotation process was performed using the free dataset annotation and management tools provided by Roboflow. This annotation process enabled the identification and separation of three classes of regions of interest (ROIs): cotyledon, hypocotyl, and root.

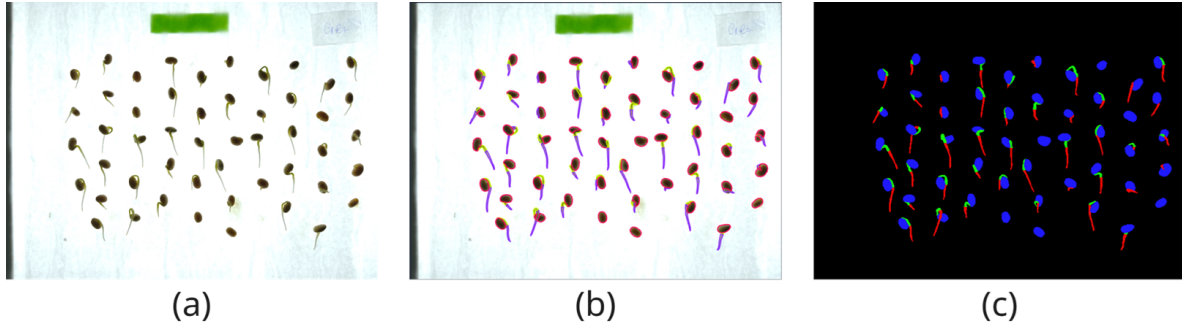


Fig 6. Examples of images during the annotation process in Roboflow. (a) Original image, (b) image with annotation polygons, and (c) generated ground truth.

After image annotation, preprocessing options were selected within the Roboflow platform to adjust the model input images during the training and validation stages. In addition, data augmentation techniques were selected and applied to the training dataset in order to increase dataset variability.

Regarding preprocessing, the following procedures were applied: (i) auto-orientation to prevent discrepancies between the locations of the annotated masks and the corresponding objects in the original images; and (ii) image resizing to 640 x 640 pixels, aiming to reduce computational cost and contribute to the optimization of training time.

With respect to the data augmentation process, horizontal and vertical flips, 90° rotations (clockwise and counterclockwise), and small random rotations of up to $\pm 10^\circ$ were applied, allowing the model to recognize patterns regardless of camera orientation or object arrangement.

Additionally, noise was added to up to 0.25% of the image pixels to improve the models' ability to handle images affected by sensor noise, dirt, or minor imperfections, which are commonly encountered in real-world scenarios. Examples of the resulting transformations are illustrated in Figure 7.

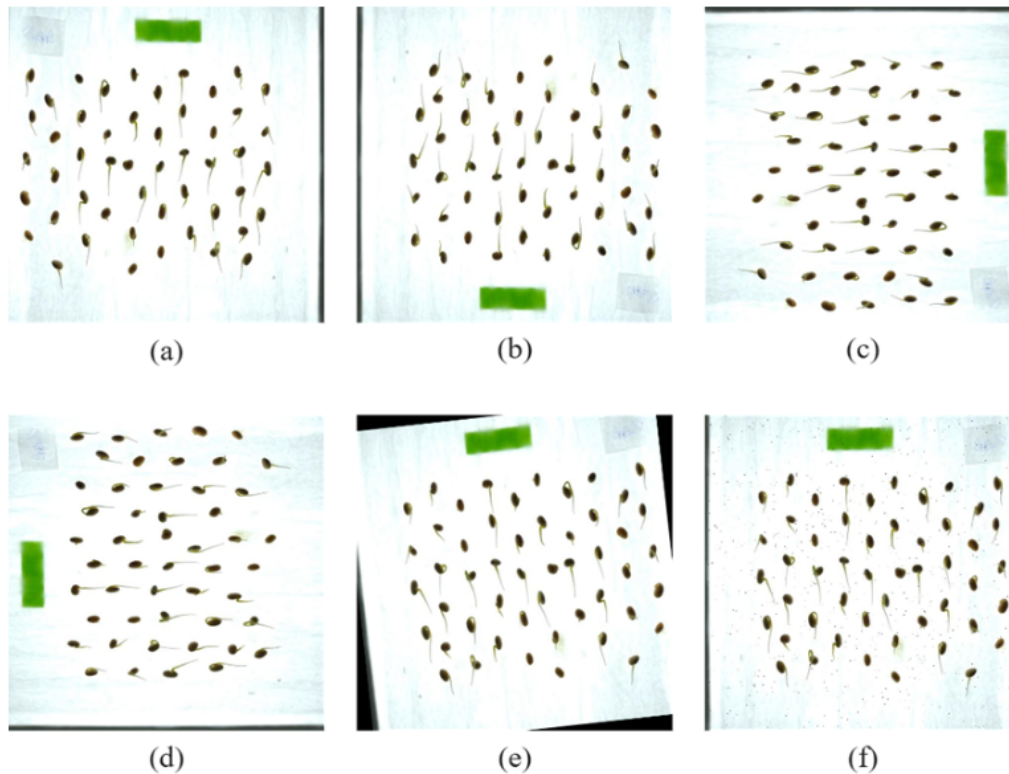


Fig 7. Examples of image transformations. (a) horizontal flip, (b) vertical flip, (c) clockwise rotation, (d) counterclockwise rotation, (e) 10° rotation, and (f) noise added to 0.25% of the pixels.

In total, the dataset consisted of 165 images for training, 16 for validation, and 8 for testing.

Experiment Design

The dataset was divided such that 70% of the images were allocated for training, 20% were reserved for validation, and the remaining 10% were used for testing. These subsets were shared among all models to ensure a fair comparison. Subsequently, synthetic images were generated from the training subset.

Most model initialization parameters were standardized during the training stage in an attempt to provide similar starting conditions for all models. The AdamW optimizer was selected, with a learning rate of 10^{-3} , a momentum value of 0.9, and a weight decay of 0.5×10^{-3} . All models were trained for 100 epochs. With regard to batch size, a value of 8 was adopted, except for the U-Net model, which, due to hardware limitations, required training with a batch size of 4.

The experiment was conducted in the Google Colab environment, a free platform that enables users to write and execute Python code through a web browser. Table 1 presents the specifications of the machine used in the development environment.

Table 1. Specifications of the machine used during the development of this study.

System and Environment	Specifications
Programming Language	Python 3.13.0
Machine Learning Frameworks	Ultralytics 8.3.49, Torch 2.8, Detectron2 0.6
Other Packages	Roboflow 1.1.49
CPU	Intel(R) Xeon(R) 2.00 GHz

GPU	Tesla T4
VRAM	15 GB
RAM	51 GB

Performance Metrics

The present study employed the metrics Recall, Precision, mAP50, and mAP50-95. Most of these metrics are derived from a confusion matrix composed of four fundamental values: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

Precision (Equation 1) is a measure of prediction accuracy, indicating the proportion of correctly predicted positive outcomes. Therefore, it is defined as the ratio between True Positives and all predicted positives, representing the proportion of pixels correctly identified as belonging to a ROI among all pixels predicted to belong to that ROI.

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

Recall (Equation 2) is the proportion of true positive cases that are correctly predicted as positive. Therefore, it measures the ability of the model to correctly identify TP instances, that is, among all pixels that truly belong to a ROI.

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

Average Precision (AP) (Equation 3) is defined as the area under the Precision–Recall curve and measures the quality of the model predictions for a single class.

$$AP = \int_0^1 Precisão(Recall)d(Recall) \quad (3)$$

Intersection over Union (IoU) (Equation 4) is a metric used to measure the overlap between the region predicted by the model and the ground-truth region.

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (4)$$

With respect to the mAP50 metric (Equation 5), the Average Precision (AP) value is calculated for all classes considering an Intersection over Union threshold of 0.50 (50%) between the model prediction and the ground truth.

$$mAP50 = \frac{1}{C} \sum_{c=1}^C AP_c \quad (5)$$

where (C) represents the total number of trained classes and (AP_c) is the Average Precision for class (c), considering an IoU threshold greater than 50%.

The mAP50-95 metric (Equation 6), in turn, corresponds to the mean Average Precision calculated over different IoU thresholds ranging from 0.50 to 0.95, with increments of 0.05 (5%), providing a comprehensive assessment of model performance across different levels of difficulty.

$$mAP50-95 = \frac{1}{C} \sum_{c=1}^C \left(\frac{1}{10} \sum_{i=1}^{10} AP_{c,i} \right) \quad (6)$$

Results and Discussion

Table 2. Average performance metrics of the Mask R-CNN model.

Class	Precision	Recall	mAP50	mAP50-95
-------	-----------	--------	-------	----------

Cotyledon	0.99	0.93	-	0.68
Hypocotyl	0.63	0.08	-	0.02
Root	0.64	0.16	-	0.04

Considering Table 2, it can be observed that the Mask R-CNN model exhibits good performance in the segmentation of relatively simple structures, such as the cotyledon. However, the Recall and mAP values for more complex structures (hypocotyl and root) are extremely low. In Figure 8(b), this relationship between Recall and segmentation quality can be clearly observed, as evidenced by the absence of hypocotyl and root segmentation in many seedlings.

Table 3. Average performance metrics of the U-Net model.

Class	Precision	Recall	mAP50	mAP50-95
Cotyledon	0.83	0.98	0.99	0.81
Hypocotyl	0.59	0.88	0.71	0.55
Root	0.68	0.89	0.77	0.63

According to the U-Net results presented in Table 3, an opposite behavior to that observed in Table 2 can be identified regarding Precision and Recall, particularly for the hypocotyl and root classes. It can be seen that the model was able to stabilize Recall values for the most complex seedling structures. A high Recall value indicates that a large proportion of pixels belonging to a given class were successfully identified. On the other hand, a reduction in Precision values can be observed, indicating that although the model effectively identifies a class, its ability to correctly classify a given pixel may be compromised. Nevertheless, U-Net achieved higher segmentation quality than the Mask R-CNN model, mainly because it was able to identify and classify the entire region of interest of the seedling, as shown in Figure 8(c).

Table 4. Average performance metrics of the YOLO-v8 model.

Class	Precision	Recall	mAP50	mAP50-95
Cotyledon	0.98	0.80	0.98	0.66
Hypocotyl	0.73	0.68	0.64	0.23
Root	0.87	0.73	0.80	0.46

Based on the data presented in Table 4, it can be observed that the segmentation results remain within a range considered satisfactory. Even so, an overall improvement in the metrics can be noted when compared with those reported in Table 3. YOLO-v8 exhibited a marked difference in the mAP50-95 metric among the analyzed classes, particularly for the hypocotyl, reinforcing the complexity of this class even in seemingly simple tasks.

According to the results in Table 4, the hypocotyl class achieved an mAP50 value of 64%, demonstrating limited improvement when compared with the results obtained by the U-Net model. In contrast, for the cotyledon and root classes, a performance increase was observed relative to U-Net, with Precision values increasing by 15% and 19%, respectively. Therefore, the hypocotyl remains the most challenging class.

The persistence of this complexity is likely attributable to two factors: (i) the low representativeness resulting from the limited size of the dataset, and (ii) the need to reduce image resolution during the training stage.

Table 5. Average performance metrics of the YOLO-v9 model.

Class	Precision	Recall	mAP50	mAP50-95
Cotyledon	0.98	0.97	0.98	0.67
Hypocotyl	0.78	0.72	0.65	0.22
Root	0.88	0.75	0.82	0.47

Upon analyzing Table 5, conclusions analogous to those drawn from Table 4 can be reached. However, improvements can be observed in the Precision, Recall, and mAP50 metrics for the hypocotyl and root classes, with the 17% increase in Recall for the cotyledon class being particularly noteworthy.

The segmentation obtained for an image from the test partition, considering all evaluated models, is presented in Figure 8. In this figure, the blue region corresponds to the cotyledon, the green

region to the hypocotyl, and the red region to the root.

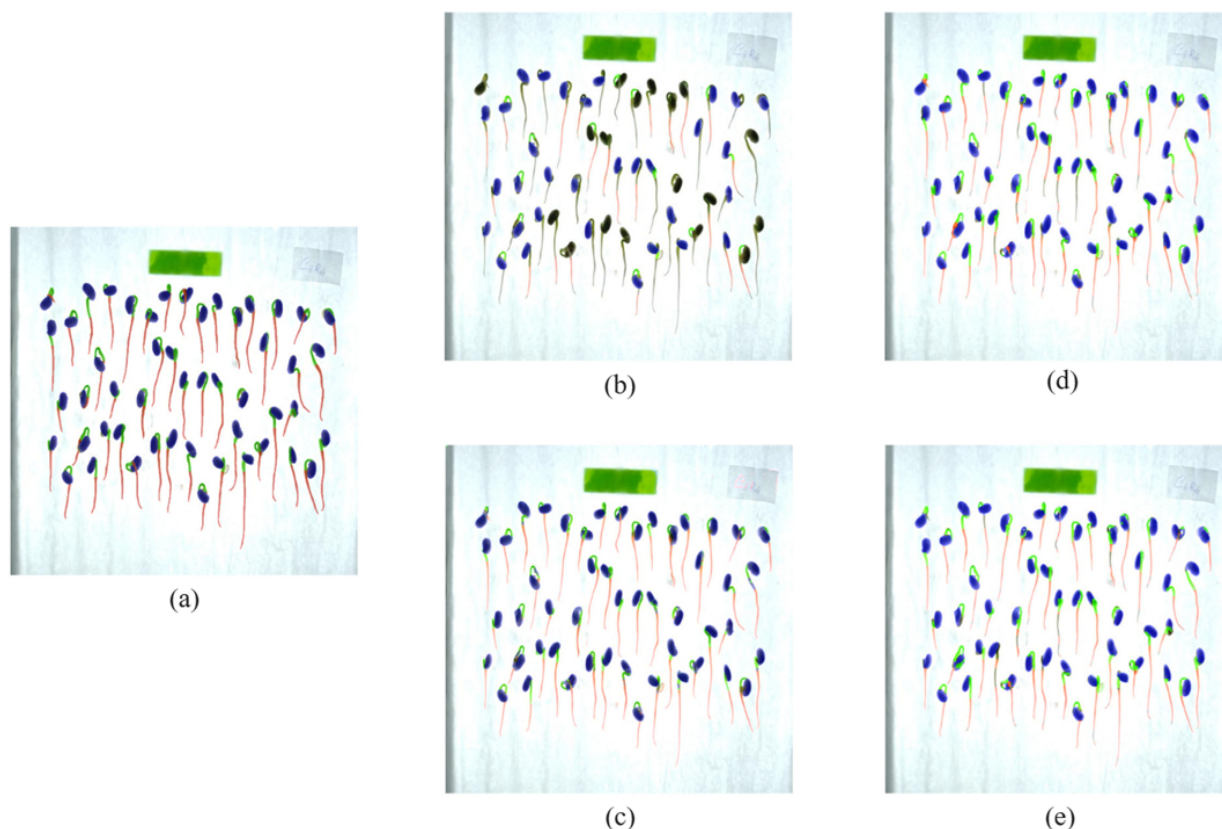


Fig 8. Results of the models on seedlings with moderate vigor. (a) labels overlaid on the original image, (b) Mask R-CNN, (c) U-Net, (d) YOLO-v8, and (e) YOLO-v9.

From a qualitative perspective, the results produced by the YOLO family models (Figures 8(d) and 8(e)) were virtually identical, with only minor localized errors being observed, which were insufficient to compromise the overall quality of the segmentations.

Therefore, based on the observations from Figure 8, the overall performance of YOLO-v9 was similar to that of YOLO-v8. However, Table 5 indicates a slight improvement in the overall segmentation results achieved by the model during the segmentation stage, leading to the conclusion that YOLO-v9 was the most suitable model for the proposed context and task.

Conclusion

Under the experimental conditions adopted in this study, the evaluated models achieved satisfactory results. Nevertheless, the high complexity of the seedling segmentation task likely contributed to limitations in the performance of all investigated models. However, models from the YOLO family, particularly YOLO-v9, demonstrated better adaptation to the application context and to the samples used during the training stage.

YOLO-v9 emerged as the best-performing model among those compared. It achieved 78% Precision and 72% Recall for the most challenging class (hypocotyl), while successfully recognizing and segmenting most of the seedling structures present in the images. Furthermore, its compact size and low inference time are factors that enable its deployment in real-world applications, further supporting its selection as the superior model. Consequently, employing a YOLO-v9-based model as a tool to replace part of the activities traditionally performed manually can streamline the analyst's workflow.

As future work, it is recommended to conduct a benchmark evaluation of the YOLO-v9 model

within a real operational environment involving human analysts in order to assess whether the reported results are satisfactorily reflected during its extensive use.

Acknowledgments

This study was supported by the Coordination for the Improvement of Higher Education Personnel (CAPES), Brazil – Finance Code 001. Danielo G. Gomes acknowledges the support of the National Council for Scientific and Technological Development (CNPq) under grant numbers 311845/2022-3 and 403996/2025-2.

References

Maja, M. M., & Ayano, S. F. (2021). The impact of population growth on natural resources and farmers' capacity to adapt to climate change in low-income countries. *Earth Systems and Environment*, 5(2), 271–283. <https://doi.org/10.1007/s41748-021-00209-6>

Kopittke, P. M., Menzies, N. W., Wang, P., McKenna, B. A., & Lombi, E. (2019). Soil and the intensification of agriculture for global food security. *Environment International*, 132, 105078. <https://doi.org/10.1016/j.envint.2019.105078>

Organização das Nações Unidas para a Alimentação e a Agricultura. (2024). Agricultural production statistics: 2010–2023. <https://www.fao.org/statistics/highlights-archive/highlights-detail/agricultural-production-statistics-2010-2023/>

Empresa Brasileira de Pesquisa Agropecuária. (2021). O agro no Brasil e no mundo. Brasília, DF. <https://www.embrapa.br/documents/10180/62618376/O+AGRO+NO+BRASIL+E+NO+MUNDO.pdf>

Smith, P., Gregory, P. J., van Vuuren, D., Obersteiner, M., Havlík, P., Rounsevell, M., Woods, J., Stehfest, E., & Bellarby, J. (2010). Competition for land. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1554), 2941–2957. <https://doi.org/10.1098/rstb.2010.0127>

Silva, D. A. da, Freitas, E. D. G. de, Abud, H. F., & Gomes, D. G. (2024). Applying YOLOv8 and X-ray morphology analysis to assess the vigor of *Brachiaria brizantha* cv. Xaraés seeds. *AgriEngineering*, 6(2), 869–880. <https://doi.org/10.3390/agriengineering6020050>

Medeiros, A. D. de, Silva, L. J. da, Ribeiro, J. P. O., Ferreira, K. C., Rosas, J. T. F., Santos, A. A., & Silva, C. B. da. (2020). Machine learning for seed quality classification: An advanced approach using merger data from FT-NIR spectroscopy and X-ray imaging. *Sensors*, 20(15), 4319. <https://doi.org/10.3390/s20154319>

Zou, Z., Chen, J., Wu, W., Luo, J., Long, T., Wu, Q., Wang, Q., Zhen, J., Zhao, Y., Wang, Y., Chen, Y., Zhou, M., & Xu, L. (2023). Detection of peanut seed vigor based on hyperspectral imaging and chemometrics. *Frontiers in Plant Science*, 14. <https://doi.org/10.3389/fpls.2023.1127108>

Filho, E. C. (2005). Desempenho de plantas oriundas de sementes de alto e baixo vigor dentro de uma população de soja (Tese de Doutorado em Ciências). Universidade Federal de Pelotas, Pelotas.

Höfs, A., Schuch, L. O. B., Peske, S. T., & Barros, A. C. S. A. (2004). Efeito da qualidade fisiológica das sementes e da densidade de semeadura sobre o rendimento de grãos e qualidade

industrial em arroz. *Revista Brasileira de Sementes*, 26(1), 55–62.

Nakagawa, J., Machado, J. R., & Rosolem, C. A. (1985). Efeito da qualidade de semente sobre o estabelecimento da população e outras características da soja. *Revista Brasileira de Sementes*, 7(2), 47–62.

Tekrony, D. M., & Egli, D. B. (1991). Relationship of seed vigor to crop yield: A review. *Crop Science*, 31, 816–822.

Vanzolini, S., & Carvalho, N. M. (2002). Efeito do vigor de sementes de soja sobre o seu desempenho em campo. *Revista Brasileira de Sementes*, 24(1), 33–41.

Dell'Aquila, A. (2009). Automated seed vigour testing: Where are we now? *Seed Science and Technology*, 37(3), 593–610.

Marcos-Filho, J. (2015). *Fisiologia de sementes de plantas cultivadas* (2^a ed.). Londrina: ABRATES.

Castan, D. O., Gomes-Junior, F. G., & Marcos-Filho, J. (2018). Vigor-S, a new system for evaluating the physiological potential of maize seeds. *Scientia Agricola*, 75(2), 167–172. <https://doi.org/10.1590/1678-992X-2016-0401>

Conrad, R. (1997). Method and apparatus for assessing the quality of a seed lot (United States Patent No. US5659623). Ball Horticultural Company.

Hoffmaster, A. L., Fujimura, K., McDonald, M. B., & Bennett, M. A. (2003). An automated system for vigour testing three-day-old soybean seedlings. *Seed Science and Technology*, 31, 701–713.

McNertney, C. (1999). System and method for measure seedlot vigor (United States Patent No. US5864984). Paradigm Research Corporation.

Gonzalez, R. C., & Woods, R. E. (2018). *Digital image processing* (4th ed.). Pearson.

Zhang, H., & Li, D. (2014). Applications of computer vision techniques to cotton foreign matter inspection: A review. *Computers and Electronics in Agriculture*, 109, 59–70. <https://doi.org/10.1016/j.compag.2014.09.004>

Anand, R., Mishra, R. K., & Khan, R. (2022). Plant diseases detection using artificial intelligence. In *Application of machine learning in agriculture* (pp. 173–190). Elsevier. <https://doi.org/10.1016/B978-0-323-90550-3.00007-2>

Xu, Y. (2024). Application of image segmentation algorithms in computer vision. *Frontiers in Computing and Intelligent Systems*, 7(3), 17–20. <https://doi.org/10.54097/gq1s6737>

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2018). Mask R-CNN. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(2), 386–397. <https://doi.org/10.1109/TPAMI.2018.2844175>

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (Lecture Notes in Computer Science, Vol. 9351, pp. 234–241). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-24574-4_28

Kadry, S., Mahajan, S., Krishna, S. S., Yellepeddi, A., & Kuppusamy, P. (2025). An automatic detection and severity levels of COVID-19 using convolutional neural network models. In *Data*

Science in the Medical Field (pp. 15–24). Springer.

Bradshaw, T. J., & McMillan, A. B. (2021). Anatomy and physiology of artificial intelligence in PET imaging. *PET Clinics*, 16(4), 471–482.

Ouguz, A., & Ertuugrul, O. F. (2023). Introduction to deep learning and diagnosis in medicine. In *Diagnostic Biomedical Signal and Image Processing Applications with Deep Learning Methods* (pp. 1–40). Elsevier.

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 580–587). Columbus, OH: IEEE. <https://doi.org/10.1109/CVPR.2014.81>

Girshick, R. (2015). Fast R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (pp. 1440–1448). Santiago, Chile: IEEE. <https://doi.org/10.1109/ICCV.2015.169>

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 779–788). Las Vegas, USA: IEEE.

Jocher, G., Chaurasia, A., & Qiu, J. (2023). Ultralytics YOLOv8 [Computer software]. GitHub. <https://github.com/ultralytics/ultralytics>

Bento, J., Paixão, T., & Alvarez, A. B. (2025). Performance evaluation of YOLOv8, YOLOv9, YOLOv10, and YOLOv11 for stamp detection in scanned documents. *Applied Sciences*, 15(6), 3154. <https://doi.org/10.3390/app15063154>

Wang, C., & Liao, H. M. (2024). YOLOv9: Learning what you want to learn using programmable gradient information. *arXiv*. <https://arxiv.org/abs/2402.13616>

Jegham, N., Koh, C. Y., Abdelatti, M., & Hendawi, A. (2024). Evaluating the evolution of YOLO (You Only Look Once) models: A comprehensive benchmark study of YOLO11 and its predecessors. *arXiv*. <https://doi.org/10.48550/arXiv.2411.00201>