Acta Scientiarum

http://www.periodicos.uem.br/ojs/
ISSN on-line: 1807-8621
Doi: 10.4025/actasciagron.v46i1.67040





# Soybean crop yield estimation using artificial intelligence techniques

**Poliana Maria da Costa Bandeira[1*], Flora Maria de Melo Villar[1], Francisco de Assis de Carvalho Pinto[1], Felipe Lopes da Silva[2] and Priscila Pascali da Costa Bandeira[1]**

[1]Departamento de Engenharia Agrícola, Universidade Federal de Viçosa, Bela Vista, 36570-000, Viçosa, Minas Gerais, Brazil. [2]Departamento de Agronomia, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil. *Author for correspondence. E-mail: poliana.bandeira@ufv.br

**ABSTRACT.** It is common to observe conventional methods for estimating soybean crop yields, making the process slow and susceptible to human error. Therefore, the objective was to develop a model based on deep learning to estimate soybean yield using digital images obtained through a smartphone. To do this, the ability of the proposed model to correctly classify pods that have different numbers of grains, count the number of pods and grains, and then estimate the soybean crop yield was analyzed. As part of the study, two types of image acquisition were performed for the same plant. Image acquisition 1 (IA1) included capturing the images of the entire plant, pods, leaves, and branches. Image acquisition 2 (IA2) included capturing the images of the pods removed from the plant and deposited in a white container. In both acquisition methods, two soybean cultivars, TMG 7063 Ipro and TMG 7363 RR, were used. In total, combining samples from both cultivars, 495 images were captured, with each image corresponding to a sample (plant) obtained through methods AI1 and AI2. With these images, the total number of pods in the entire dataset was 46,385 pods. For the training and validation of the model, the data was divided into subsets of training, validation, and testing, representing, respectively, 80, 10, and 10% of the total dataset. In general, when using the data from IA2, the model presented errors of 7.50 and 5.32% for pods and grains, respectively. These values are considerably lower than when the model used the IA1 data, where it presented errors of 34.69 and 35.25% for pod and grain counts, respectively. Therefore, the data used from IA2 provide better results to the model.

**Keywords:** deep learning; image acquisition; smartphone; grain and pod count.

Received on February 8, 2023.
Accepted on September 14, 2023.

## Introduction

Commodity productivity estimation, such as the soybean crop, has become a strategic issue for agro-industrial companies, farmers, and the government, as it directly impact the country's economy (Ramos, Prieto, Montoya, & Oliveros, 2017). Productivity estimation is important due to its role in food management (Terliksiz & Altylar, 2019), establishing definitions of future prices for agricultural productions.

It is natural to use conventional methods to estimate soybean yield. One of these methods is based on the manual counting of pods and grains, which can be performed after pod-filling. Another method is based on the weight of grains collected from a harvested area. However, the yield estimate should be made at the R8 reproductive stage in this case. This stage corresponds to the end of the soybean's growth cycle, which is close to the harvest season. Consequently, weighing the grains at this phase allows for a more accurate estimation of the expected yield. These methods have limitations such as human error and may also create difficulties as the process is carried out close to the harvest.

In recent years, unconventional methods for performing yield estimation have emerged, using artificial neural networks and convolutional neural networks. These approaches apply all deep learning technology in the process of yield estimation in agricultural crops (Alves, Teixeira, Melo, Souza, & Silva, 2018; Chan, Wei, & Molin, 2020; Miller et al., 2018; Sun, Di, Sun, Shen, & Lai, 2019; Terliksiz & Altylar, 2019). Despite being efficient solutions, these methods have some limitations, such as increased costs for the producer when compared to the conventional methods and the need to have data from the previous crops.

As mentioned, the newer methods for soybean yield estimation use deep learning models, which can extract several inherent features of the object of interest, thus enabling object recognition and classification. This process is based on convolutional neural networks (CNNs) present in deep learning models. CNNs have emerged as promising tools for pattern recognition in images since they can produce filters that help predict classes associated with the training images (Maxwell, Pourmohammadi, & Poyner, 2020). The efficiency of this technique made it popular for applications in the agricultural sector (Apolo-Apolo, Martínez-Guanter, Egea, Raja, & Pérez-Ruiz, 2020; Arsenovic, Karanovic, Sladojevic, Anderla, & Stefanovic, 2019; Durmus, Gunes, & Kirci, 2017; Ganesh, Volle, Burks, & Mehta, 2019).

The Mask R-CNN model has been highlighted among various deep learning models. This model shows high efficiency in instance segmentation (He, Gkioxari, Dollár, & Girshick, 2020). The *Mask R-CNN* algorithm is an extension of the *Faster R-CNN* model (Ren, He, Girshick, & Sun, 2015). However, *Mask R-CNN* has a mask prediction branch. This branch is a fully convolutional network that is applied in each region of interest, and thus, the segmentation mask is generated to execute the object classification. Thus, the *Mask R-CNN* model has shown its potential for use in agricultural processes (Carvalho et al., 2021; Lee, Nazki, Baek, Hong, & Lee, 2020; Mekhalfi, Nicolò, Bazi, Al Rahhal, & Al Maghayreh, 2021).

Deep learning models use digital images as essential data. One source of obtaining this type of data is the smartphone, which can be easily used as it is owned by the general population. In the agricultural field, smartphones are already being extensively used in various processes as a tool to collect data to estimate cotton yields, as described by Tedesco-Oliveira, Pereira, Maldonado Jr., and Zerbato (2020), or as a resource to help identify tomato leaf diseases, as mentioned by Ngugi, Abelwahab, and Abo-Zahhad (2020). Thus, this work aimed to develop a model based on deep learning, to estimate the soybean yield in the field using digital images obtained by a smartphone.

## Material and methods

Two soybean cultivars were used in this work, TMG 7063 IPRO (cultivar 1) and TMG 7363 RR (cultivar 2). The cultivars were sown on 12/02/2020 in the experimental area of the *Universidade Federal de Viçosa* (UFV) in Viçosa, Minas Gerais State, Brazil. For both cultivars, the inter-row spacing was 0.5 m, with an average of 13 and 15 plants per linear meter for cultivar 1 and cultivar 2, respectively.

To develop a model to estimate the soybean crop productivity from digital images obtained by a smartphone, the following steps were necessary: i) Construction of the database, ii) Pre-processing of the obtained images, iii) Training of the deep learning model, iv) Estimate calculation of soybean yield estimation, and v) Model evaluation.

### Database construction

Digital images for the database were acquired 77 days after sowing, between February 18 and March 04, 2021, with the soybean plants in the phenological stages between R6 and R8. The images were taken in the field from 8:00 am to 12:00 pm. Thus, we ensured that the images had different levels of brightness.

The same smartphone was used throughout the image acquisition process, which had a dual camera of 48 and 5 megapixels with a 1/2" sensor and lens opening of 1.8. All other camera settings were kept on automatic during the database construction. According to the features of the smartphone used, the generated images had a size of 3000 × 4000 pixels, and their output was in JPG format.

Two image acquisition methods were used for acquiring images of the plant with leaves, branches, and pods (method IA1) and images of the pods isolated from the plant (method IA2). For IA1, the soybean plants were removed from the experimental area and placed in a structure with a monochromatic background (Figure 1a). The plants were chosen in such a way as to include plants of large, medium, and small sizes, ensuring the training of the model plants of various sizes. The structure (Figure 1b) had a place to position the smartphone and the plant, ensuring a fixed distance of one meter between the plant and the smartphone during image acquisition.

For the IA2 method, the pods were removed from the plant and placed on a white tray (Figure 2) to capture the images. Thus, the pods of a single plant were placed on the tray at a time. During the images acquisitions for IA2, there was no standardization of the distance between the smartphone camera and the tray. For the same plant, the images were first captured as per IA1 and, then, according to IA2, created a database for each acquisition method.

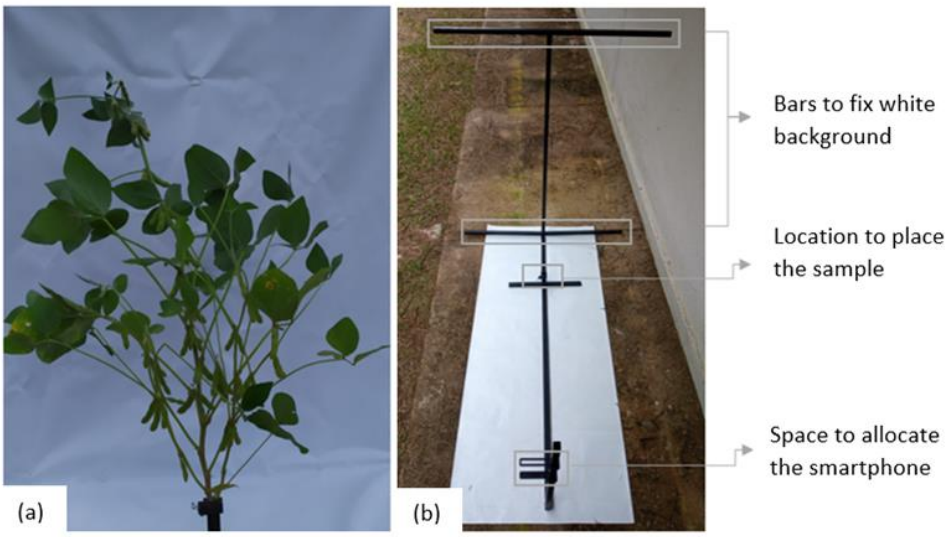



**Figure 1.** Image captured from pods of a single plant with image acquisition method 1 (IA1) (a). The frame used for sample positioning with image acquisition method 1 (IA1) (b).

**Figure 2.** Image captured from single plant pods with image acquisition method 2 (IA2).

For each cultivar, 495 plants from the experimental area were used. Each plant generated an image for IA1 and IA2. From the total number of images, considering the two cultivars and the two forms of image acquisition, the databases had a total of 46,385 pods.

## Pre-processing

Each database was randomly subdivided into sets of training images (80%), tests (10%), and validation (10%). After dividing the sets, the data was amplified, improving the algorithm performance since it artificially generates more images for training (Gonzalez, 2007). The processing steps performed to generate artificial images involved random rotations of the images at angles of 0°, 90°, 180°, and 270°, thereby altering the position of the objects of interest. Additionally, an enhancement process was applied to sharpen the edges, providing a more detailed perception. However, this method also introduces some noise, creating a challenging environment for the model. Moreover, a Gaussian filter, acting as a low-pass filter, was also used to attenuate higher frequencies in the images, resulting in a smoothing effect. Other modifications included vertical and horizontal mirroring of the images, as well as adjustments to their contrast. Each process was conducted by randomly selecting images from the training dataset. As a result, each image underwent only one of the mentioned processes, ensuring diversity in the artificially generated images.

Before performing the model training, the regions of interest (pods) were manually drawn on all subset images using the VGG *Image Annotator* online software (Dutta & Zisserman, 2019). For each region of interest, a class was given and named according to the number of grains the pod had. Thus, we had the classes "one", "two", "three" and "four", representing the pods that had up to four grains. Figure 3 and Figure 4 shows an example of the masks generated from the labeled data for IA1 and IA2, respectively.

**Figure 3.** Example of the generated mask (right) after labeling the original image (left) obtained through IA1.

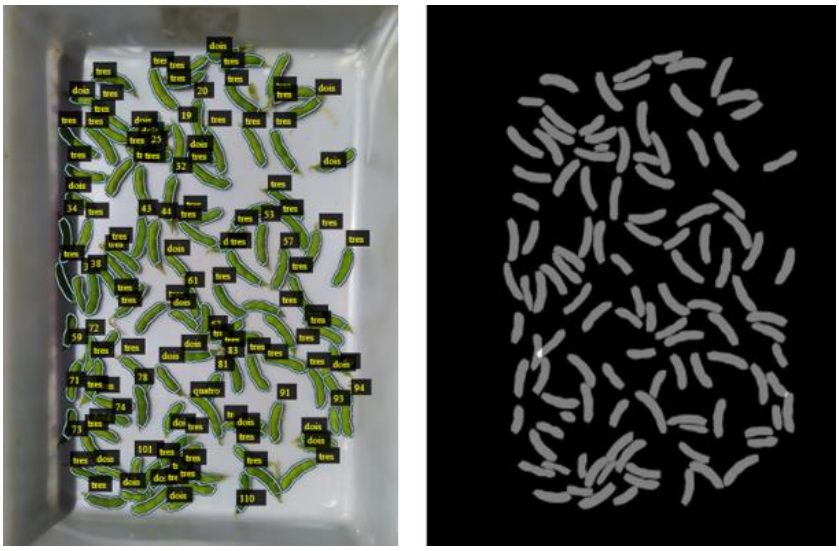


**Figure 4.** Example of the generated mask (right) after labeling the original image (left) obtained through IA2.

## Model training

The model's training was based on *Mask R-CNN* and was executed separately for the IA1 and IA2 databases. Two categories of input data, the training dataset (training and validation images) and the training data labels (training and validation labels) were required to train the model. Residual backbone network Resnet-101 was used for model training in both acquisition methods using a *Feature Pyramid Network* (FPN) type of deep neural network. Resnet-101 has an architecture with 101 layers of depth and can classify over 1,000 classes (Ghosal et al., 2019). The entire process was carried out using the Python programming language.

## Yield estimation

To estimate soybean yield, information from the two cultivars studied in this work was necessary, such as the average number of pods per plant, average number of grains per pod, number of soybean plants per hectare, and the 100-grain mass. The average number of pods per plant and the average number of grains per pod were obtained from the model developed in this work, which was used for counting. In order to validate the model, manual counting of the number of grains and pods was performed.

The number of plants per hectare was calculated by counting the plants in a known area and then calculating for one hectare. The weighing of the 100-grain mass occurred after measuring their humidity. Then the weight was corrected (Equation 1) for the humidity of 13%. The 100-grain mass and the number of plants per hectare were determined for both cultivars. The estimation of soybean productivity for each cultivar was performed according to Equation 2.

$$MC\ (g) = MU \times \frac{100-UA\ (\%)}{100-UD\ (\%)} \quad (1)$$

where *MC* is the mass corrected for desired humidity, *MU* is the grain mass for the current humidity, and both *MC* and *MU* are in grams. The variables *UA* and *UD* refer to the current and desired humidity, respectively, in percentage.

$$prod\ (sc\ ha^{-1}) = \frac{\frac{plt}{ha} \times \frac{pod}{plt} \times \frac{grains}{pod} * MC\ (g)}{6000} \quad (2)$$

where *prod* refers to the crop productivity, *plt ha*$^{-1}$ is the number of plants per hectare, *pod plt*$^{-1}$ is the average number of pods per plant, and *grains pod*$^{-1}$ is the average number of grains per pod.

### Model evaluation

### Classification

It is worth noting that the data utilized in this analysis, due to their authentic nature and their origin from field surveys, exhibit an unbalanced distribution. Methods are currently available to address the class imbalance. Nevertheless, owing to the substantial scale of the dataset (comprising 46,385 pods), implementing these techniques presents an additional challenge. As an alternative, precision, recall, and F1-Score metrics were adopted. These metrics prove to be more suitable for evaluating model performance within an unbalanced class context, as they offer a more comprehensive understanding of how the model manages distinct classes, enabling a meticulous analysis of outcomes in terms of false positives, false negatives, and accurately classified examples.

From the relationship between the real classification, based on manual labeling, and that predicted by the model, it was possible to calculate the metrics of accuracy (Equation 3), recall (Equation 4), and F1 score (Equation 5) to evaluate the algorithm's classification performance. Precision measures the proportion of predicted values that are real values. Recall measures the proportion of real values that the model could correctly classify. The F1 score is the harmonic mean between the two values, accuracy and recall, i.e., a single value to evaluate the classification.

$$P = \frac{TP}{TP+FP} \quad (3)$$

$$R = \frac{TP}{TP+FN} \quad (4)$$

$$F1 = 2 \times \frac{P \times R}{P+R} \quad (5)$$

where *P* refers to precision, *R* refers to recall, *F1* refers to F1 score, *TP* are True Positive values, *FP* are False Positive values, and *FN* are False Negative values.

True Positive (TP) refers to the model prediction coinciding positively with the accurate prediction. False Positive (FP) occurs when the object is not of interest but is predicted as an object of interest. False Negative (FN) occurs when there is an object of interest in the image, but it is not detected or is incorrectly classified.

### Counting the number of pods and grains and estimating yield

To evaluate the difference between the manually counted values and those predicted by the model, we used the mean absolute percentage error (MAPE) metric, calculated according to Equation 6. This metric is commonly used to evaluate productivity estimation (Nevavuori, Narra, & Lipping, 2019; Tedesco-Oliveira et al., 2020).

$$MAPE = \frac{100}{n} \times \sum_{t=1}^{n} \left| \frac{C_{real} - C_{predicted}}{C_{real}} \right| \quad (6)$$

where *n* is the total number of samples, $C_{real}$ and $C_{predicted}$ are the real and predicted values for yield estimation and grain and pod count, respectively.

## Results

### Object classification and detection

The pods that were classified by the model in the test set, for methods IA1 and IA2, were distributed in a heat map (Figure 5). Thus, it is possible to identify how many objects were correctly classified, how many were

wrongly classified, and how many could not be detected by the model. The points in red denote greater intensity, it is noted that the data classified by the IA1 method presented a higher value of undetected objects when compared with the data from IA2. In addition, it is observed that correct classifications were more for data from IA2 when compared to IA1.


**Figure 5.** Heat map for pods classified by the model in the test set of IA1 (a) and IA2 (b) methods.

The evaluation for IA1 and IA2 from the accuracy, recall, and F1 score metrics are presented in Figure 6a, b, and c, respectively. The model presented better results for IA2 than IA1. This may be because the objects were covered in IA1, thus hindering the correct identification. Furthermore, among the four classes analyzed, it is observed that in classes "two" and "three," the model presented a better performance for both methods of image acquisition.

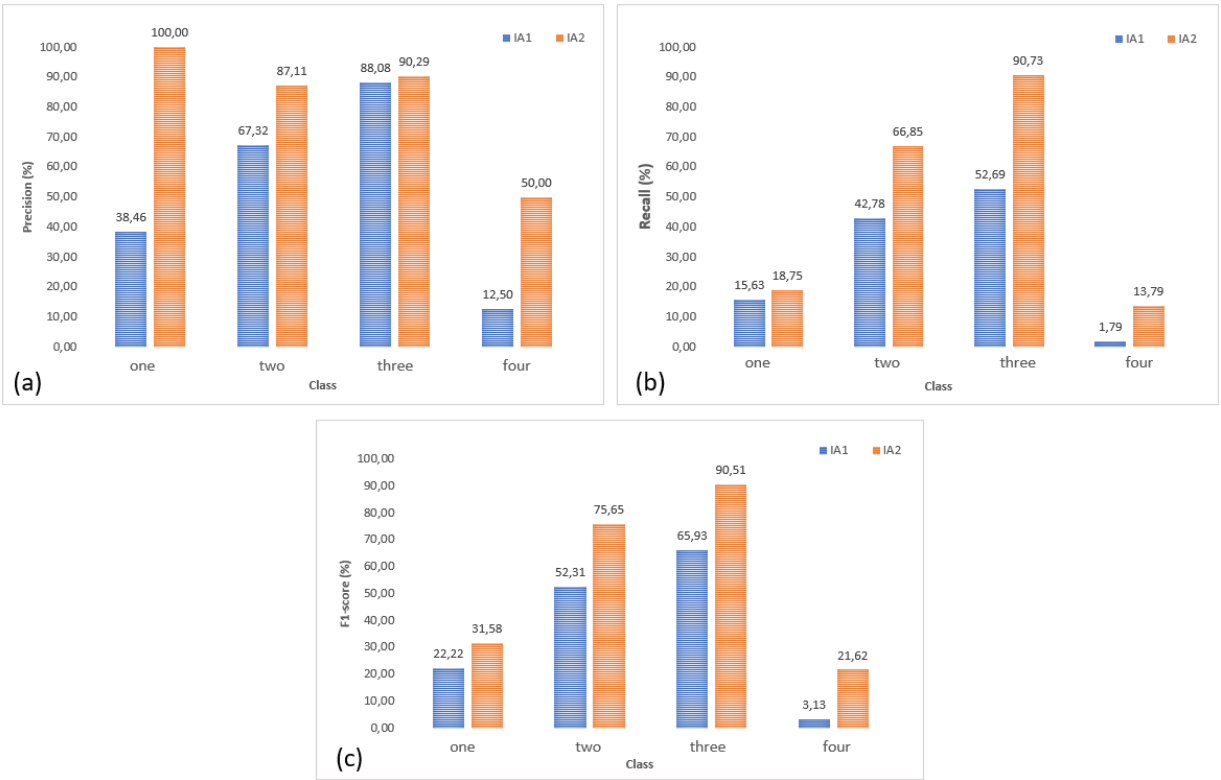



**Figure 6.** Model classification performance based on the accuracy (a), recall (b), and F1 score (c) metrics for both image acquisition methods: image acquisition 1 (IA1) and image acquisition 2 (IA2).

Different situations were observed in the data classification performed by the model by using the IA1 method. Figure 7a shows a difficulty imposed by the data, the obstruction of the pod by the branch, however, the model overcame this difficulty and correctly classified the pod in the image. This result was not repeated in Figure 7b, where the model failed to classify, the presence of noise (leaves and branches) in the IA1 data probably caused this type of failure, which led to incorrect identification.


**Figure 7.** Different situations for model detection using IA1 data: pod obstruction caused by noise (leaves and branches) (a) and erroneous detection caused by noise (leaves and branches) (b).

## Evaluation of pod and grain count

Figure 8 shows, for the test set of both acquisition methods, the manually counted values and those predicted by the model referring to the number of pods and grains. The MAPE for the set referring to the IA1 method was found to be 34.69 and 35.25% for pods and grains, respectively. In contrast, the model using the IA2 data had a MAPE of 7.50 and 5.32% for pods and grains across the test set, respectively.

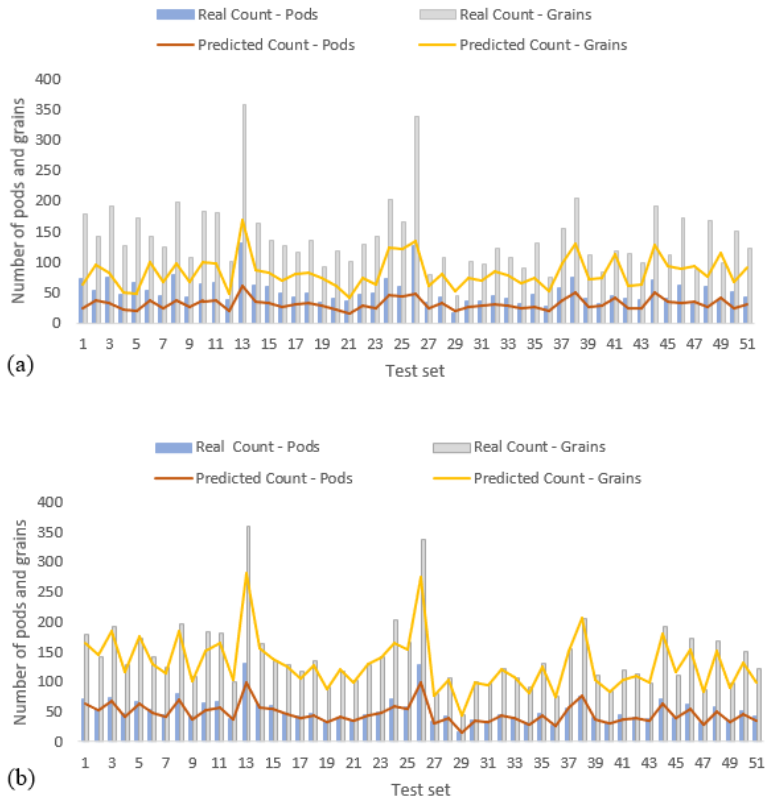


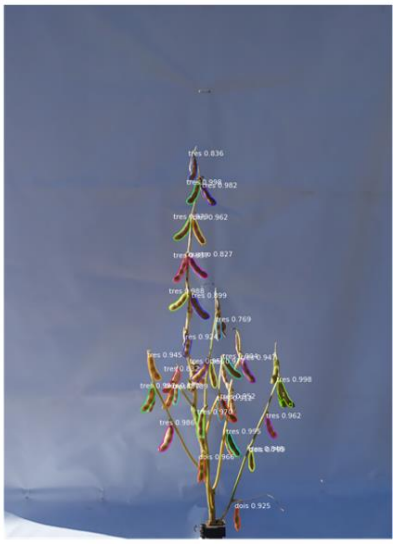

**Figure 8.** Manual and model predicted counts for the number of pods and grains with both image acquisition methods: image acquisition 1 (IA1) and image acquisition 2 (IA2).

According to the data generated by the IA1 method, it was observed that some images presented a large difference between the manual count and that predicted by the model for the number of pods and grains, while other images presented a small difference. Figure 9a shows a plant in the reproductive stage R6, which has a large number of leaves and branches and also showed a more significant error. Figure 9b shows a plant in the R8 stage with fewer leaves and branches and presents a smaller error. This confirms the negative impact of leaves and branches in classifying pods based on the number of grains they have.

(a)

| | Pods | Grains |
|---|---|---|
| Real Values | 47 | 128 |
| Predicted Values | 21 | 51 |
| MAPE (%) | 55.32 | 60.16 |

(b)

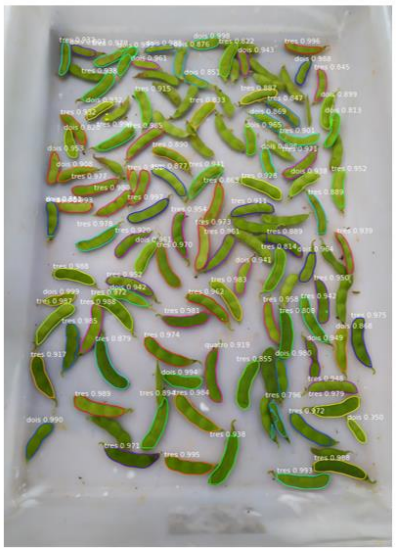


| | Pods | Grains |
|---|---|---|
| Real Values | 42 | 123 |
| Predicted Values | 31 | 91 |
| MAPE (%) | 26.19 | 26.02 |

**Figure 9.** Calculation of the ASM is influenced by the plant's reproductive stage: plant in R6 with the greatest number of leaves and branches (a). Plant in R8 with fewer leaves and branches (b).

It is observed that the number of pods inside the tray in the IA2 data contributed to the increase or decrease in the error for pod and grain counting. Figure10a shows that a greater quantity of pods inside the tray led to a high error, while in Figure 10b, as the quantity of pods is reduced, the error between the real and predicted values is also reduced.

(a)



| | Pods | Grains |
|---|---|---|
| Real Values | 128 | 339 |
| Predicted Values | 100 | 279 |
| MAPE (%) | 21.88 | 18.58 |

(b)

| | Pods | Grains |
|---|---|---|
| Real Values | 44 | 122 |
| Predicted Values | 43 | 120 |
| MAPE (%) | 2.27 | 1.64 |

**Figure 10.** Tray with the largest number of pods causing a greater error in the model classification (a) and with a smaller number of pods generating a smaller error in the model classification (b).

## Productivity estimation

Table 1 shows the error generated and the difference between the estimate calculations performed using the conventional method and the methods discussed in this article for both cultivars. Contrasting values were

observed when the model used IA1 and IA2. In the case of IA2, the estimated productivity by the model was close to the estimate calculated by the conventional method. In contrast, in the case of IA1, a large difference was observed between the prediction of the model and that of the conventional method for estimating soybean productivity.

**Table 1.** Mean absolute percentage error (MAPE) between yield estimates with the conventional method and predicted by the model for the two soybean cultivars.

| Soybean cultivars | MAPE IA1 (%) | MAPE IA2 (%) |
|---|---|---|
| Cultivar 1 | 44.76 | 6.29 |
| Cultivar 2 | 32.73 | 5.50 |

## Discussion

In this study, a model that is capable of classifying soybean pods with a different number of grains was developed. Based on the pod classification, the model was used to count the number of pods and grains, and, from this, the soybean crop yield was estimated. The algorithm was developed for use in the field, so during the database construction, the images captured were subject to changing environmental conditions.

The classification of the pods using the Mask R-CNN model has been presented in the heat maps (Figure 5). The values refer to the pods present in the images of the test subset for IA1 and IA2, and the model was used for classification. Based on the heat maps, a large number of undetected objects is noticeable when the model uses IA1. In contrast, when IA2 data was used, the model presented a superior performance in detection, besides showing a greater accuracy of classification. This caused the model to perform better on the classification metrics using the IA2 data (Figure 6), in which the metrics presented the highest values.

In general, it was observed that, regardless of the procedure used to obtain the images, the model performed better in the classification of classes "two" and "three" than that of classes "one" and "four". This result is possibly due to the dominance of classes "two" and "three" in the entire database, which reached 96.06% compared to classes "one" and "four". There were more samples for classes "two" and "three" and fewer samples for classes "one" and "four" at the time of model training. This probably caused the disparity in the performances between the classes, as presented in Figure 6.

It is worth mentioning that obtaining data in the way described in the work and that are balanced is something complex and difficult to achieve. Therefore, the metrics used allow us to identify and better understand the results obtained. In future work, it is suggested to explore data pre-processing techniques to deal with imbalance, such as oversampling or undersampling, in order to improve the representation of minority classes and thus contribute to even more robust and reliable results.

It was observed that the accuracy for class "one" achieved by the model using the IA2 data reached a value of 100% (Figure 6a). This metric indicates the percentage of predictions that correspond to the real value. Thus, even though not all objects of this class were correctly predicted, the model could classify class "one" correctly. However, the opposite occurs when analyzing recall, with a value of 18.75% (Figure 6b). This metric indicates the percentage of real values that were correctly predicted. Furthermore, as observed, the model failed in some cases, classifying incorrectly or not detecting all the pods of class "one".

Despite the low performance of the model in the classification using IA1 data compared to the performance using the IA2 images, it was observed that the IA1 images presented a higher number of objects that are not of interest, such as leaves and branches. Even so, in certain instances, the model overcame these difficulties imposed by the IA1 data and located the pods that had an obstruction and classified them correctly based on the number of grains they had (Figure 7a). However, when using the IA1 data, the model sometimes misclassified the pods by detecting background objects as pods (Figure 7b). As already mentioned, the IA1 images had a greater number of objects in the background when compared with the IA2 images, causing this difficulty and confusion for the model, leading to a greater propensity for error.

The Mask R-CNN model has presented promising results in recent works, for object detection and counting in images. For example, Xu et al. (2020) used *Mask R-CNN* to count cattle in the pastures, with error values between 6 and 10%. In another study, Davis et al. (2020) used the Mask R-CNN model, identified and counted several species of herbs from phenological data of shoots, flowers, and fruits, reaching coefficient of determination ($R^2$) values of 0.71, considering all the species used.

Specifically for soybean, researchers have developed methods to perform pod and grain number counting using recent technologies. However, these solutions have some limitations. In the case of Uzal et al. (2018) e Li et al. (2019), in which both used CNNs, an accuracy of 86.20% and a mean absolute error of 13.2 were achieved. When compared with the results presented in this paper, both articles achieved good performances with their respective models. However, the experiments were performed in an environment with controlled light, making it impossible to use them in the field.

In the present study, the developed model achieved for the IA1 method, a MAPE of 34.64 and 35.25% for pods and grains, respectively. These values show the low performance of the model when using the IA1 data to perform the pod and grain counts. In contrast, the results obtained with IA2 were significantly better; the MAPE value was 7.50 and 5.32% for pods and grains, respectively. These values presented by the model when using the data obtained by IA2 prove the model's efficiency to count soybean pods and grains.

In Figure 8a, we can observe images of the test subset of the IA1 data that showed a high disparity between the values of the manual count and those predicted by the model, both for pods and grains. The difference between the images that have contrasting errors is because the plants were in the reproductive stages. The data that presented the greatest disparities between the real and predicted values for counting the pods and grains number are images of plants in the reproductive stage R6 when the soybean plant has abundant leaves and branches (Figure 9a). Those that presented a smaller disparity were plants in the R8 reproductive stage, thus presenting few leaves and branches (Figure 9b).

The difference between the real and predicted count depends on the reproductive stages; this reinforces how the branches and leaves can interfere negatively in the detection and counting of the pods and grains number that the model can perform on the IA1 data. However, it is necessary to conduct a better analysis, generating a model individually for each reproductive stage and determining the progression of the model performance in the reproductive stages with a greater number of leaves and branches than those with fewer.

Despite the model's good performance when using IA2 to count pods and grains, it was observed that the trays with the greatest number of pods caused errors (Figure 10). This may have occurred due to the presence of overlapping pods, making it difficult for the model to identify and classify the pods below other pods. In contrast, this was observed to a lesser extent in trays that had fewer pods as there was more space between the pods (Figure 10b). Thus, at the time of image acquisition, the model's performance can be improved by positioning the pods in a container with larger dimensions, providing more space between pods.

Finally, it was possible to estimate the yield of the two soybean cultivars using the data obtained by the model. Again, IA2 presented a better result when compared to IA1 for the estimation of soybean yields (Table 1). The average absolute percentage errors were, 6.29 and 5.50% for IA2 for cultivars 1 and 2, respectively, while IA1 had errors of 76 and 32.73% for cultivars 1 and 2, respectively.

The results presented by the model when using the data obtained by IA2 open up a possibility for future work. For example, an application can be developed for mobile devices as they are inexpensive and straightforward tools. With the creation of an application for mobile devices, the estimation of soybean productivity would become an accessible process that can be done by researchers and farmers.

## Conclusion

A model based on Mask R-CNN was developed to perform soybean yield estimation using smartphone images, through two image acquisition methods, which were named AI1 (image acquisition 1) and AI2 (image acquisition 2). The model classified the pods based on the number of grains they presented and was able to count the number of pods and grains and finally estimate soybean yield. For the two image acquisition methods used, the model achieved the best performance in the pod classification, pod and grain number counting, and yield estimation tasks when using the data obtained by the IA2 method. The error concerning yield estimation when the model used the IA2 data reached 6.29 and 5.50% for the cultivars analyzed, while for the IA1 data an error of 44.76 and 32.73% was presented. Therefore, using the data from the IA2 method, the developed model becomes a viable and efficient tool to estimate soybean yield in the field. The system developed in this work can be useful for the development of a commercial application for mobile devices, making it easy and fast to estimate soybean yields in the field.

## Acknowledgements

The authors of this article thank CNPq (National Council for Scientific and Technological Development) and CAPES (Coordination for the Improvement of Higher Education Personnel) for the financial support in this work.

## References

Alves, G. R., Teixeira, I. R., Melo, F. R., Souza, R. T. G., & Silva, A. G. (2018). Estimativa da produtividade de soja com redes neurais artificiais. *Acta Scientiarum. Agronomy*, *40*(1), 1-9. DOI: https://doi.org/10.4025/actasciagron.v40i1.35250

Apolo-Apolo, O. E., Martínez-Guanter, J., Egea, G., Raja, P., & Pérez-Ruiz, M. (2020). Deep learning techniques for estimation of the yield and size of citrus fruits using a UAV. *European Journal of Agronomy*, *115*, 126030. DOI: https://doi.org/10.1016/j.eja.2020.126030

Arsenovic, M., Karanovic, M., Sladojevic, S., Anderla, A., & Stefanovic, D. (2019). Solving current limitations of deep learning based approaches for plant disease detection. *Symmetry*, *11*(7), 1-21. DOI: https://doi.org/10.3390/sym11070939

Carvalho, O. L. F., Carvalho Júnior, O. A., Albuquerque, A. O., Bem, P. P., Silva, C. R., Ferreira, P. H. G., ... Borges, D. L. (2021). Instance segmentation for large, multi-channel remote sensing imagery using mask-RCNN and a mosaicking approach. *Remote Sensing*, *13*(1), 1-24. DOI: https://doi.org/10.3390/rs13010039

Davis, C. C., Champ, J., Park, D. S., Breckheimer, I., Lyra, G. M., Xie, J., ... Bonnet, P. (2020). A new method for counting reproductive structures in digitized herbarium specimens using mask R-CNN. *Frontiers in Plant Science*, *11*, 1–13. DOI: https://doi.org/10.3389/fpls.2020.01129

Durmus, H., Gunes, E. O., & Kirci, M. (2017). Disease detection on the leaves of the tomato plants by using deep learning. In *6th International Conference on Agro-Geoinformatics, Agro-Geoinformatics 2017*. Fairfax, VA: IEEE. DOI: https://doi.org/10.1109/Agro-Geoinformatics.2017.8047016

Dutta, A., & Zisserman, A. (2019). The VIA annotation software for images, audio and video. In *MM 2019 - Proceedings of the 27th ACM International Conference on Multimedia* (p. 2276-2279). New York, NY: Association for Computing Machinery. DOI: https://doi.org/10.1145/3343031.3350535

Ganesh, P., Volle, K., Burks, T. F., & Mehta, S. S. (2019). Deep orange: Mask R-CNN based orange detection and segmentation. *IFAC-PapersOnLine*, *52*(30), 70-75. DOI: https://doi.org/10.1016/j.ifacol.2019.12.499

Ghosal, P., Nandanwar, L., Kanchan, S., Bhadra, A., Chakraborty, J., & Nandi, D. (2019). Brain tumor classification using ResNet-101 based squeeze and excitation deep neural network. In *2nd International Conference on Advanced Computational and Communication Paradigms, ICACCP 2019*, *May 2020*. Gangtok, IN: IEEE. DOI: https://doi.org/10.1109/ICACCP.2019.8882973

Gonzalez, T. F. (2007). *Handbook of approximation algorithms and metaheuristics*. New York, NY: Chapman and Hall/CRC Press. DOI: https://doi.org/10.1201/9781420010749

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2020). Mask R-CNN. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *42*(2), 386-397. DOI: https://doi.org/10.1109/TPAMI.2018.2844175

Lee, J., Nazki, H., Baek, J., Hong, Y., & Lee, M. (2020). Artificial intelligence approach for tomato detection and mass estimation in precision agriculture. *Sustainability*, *12*(21), 1-15. DOI: https://doi.org/10.3390/su12219138

Li, Y. U. E., Jia, J., Zhang, L. I., Khattak, A. M., Sun, S. H. I., Gao, W., & Wang, M. (2019). Soybean seed counting based on pod image using two-column convolution neural network. *IEEEAccess*, *7*, 64177-64185. DOI: https://doi.org/10.1109/ACCESS.2019.2916931

Maxwell, A. E., Pourmohammadi, P., & Poyner, J. D. (2020). Mapping the topographic features of mining-related valley fills using mask R-CNN deep learning and digital elevation data. *Remote Sensing*, *12*(3), 1-23. DOI: https://doi.org/10.3390/rs12030547

Mekhalfi, M. L., Nicolò, C., Bazi, Y., Al Rahhal, M. M., & Al Maghayreh, E. (2021). Detecting crop circles in google earth images with mask R-CNN and YOLOv3. *Applied Sciences*, *11*(5), 1-12. DOI: https://doi.org/10.3390/app11052238

Miller, J. J., Schepers, J. S., Shapiro, C. A., Arneson, N. J., Eskridge, K. M., Oliveira, M. C., & Giesler, L. J. (2018). Characterizing soybean vigor and productivity using multiple crop canopy sensor readings. *Field*

*Crops Research*, *216*, 22-31. DOI: https://doi.org/10.1016/j.fcr.2017.11.006

Nevavuori, P., Narra, N., & Lipping, T. (2019). Crop yield prediction with deep convolutional neural networks. *Computers and Electronics in Agriculture*, *163*, 104859. DOI: https://doi.org/10.1016/j.compag.2019.104859

Ngugi, L. C., Abelwahab, M., & Abo-Zahhad, M. (2020). Tomato leaf segmentation algorithms for mobile phone applications using deep learning. *Computers and Electronics in Agriculture*, *178*, 105788. DOI: https://doi.org/10.1016/j.compag.2020.105788

Ramos, P. J., Prieto, F. A., Montoya, E. C., & Oliveros, C. E. (2017). Automatic fruit count on coffee branches using computer vision. *Computers and Electronics in Agriculture*, *137*, 9-22. DOI: https://doi.org/10.1016/j.compag.2017.03.010

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN : towards real-time object detection with region proposal networks. *arXiv:1506.01497 [cs.CV]*, *1*, 1-14. DOI: https://doi.org/10.48550/arXiv.1506.01497

Sun, J., Di, L., Sun, Z., Shen, Y., & Lai, Z. (2019). County-level soybean yield prediction using deep CNN-LSTM model. *Sensors*, *19*(20), 1-21. DOI: https://doi.org/10.3390/s19204363

Tedesco-Oliveira, D., Pereira, R., Maldonado Jr., W., & Zerbato, C. (2020). Convolutional neural networks in predicting cotton yield from images of commercial fi elds. *Computers and Electronics in Agriculture*, *171*, 105307. DOI: https://doi.org/10.1016/j.compag.2020.105307

Terliksiz, A. S., & Altylar, D. T. (2019). Use of deep neural networks for crop yield prediction: A case study of soybean yield in lauderdale county, Alabama, USA. In *8th International Conference on Agro-Geoinformatics, Agro-Geoinformatics 2019* (p. 9-12). Istanbul, TR: IEEE. DOI: https://doi.org/10.1109/Agro-Geoinformatics.2019.8820257

Uzal, L. C., Grinblat, G. L., Namías, R., Larese, M. G., Bianchi, J. S., Morandi, E. N., & Granitto, P. M. (2018). Seed-per-pod estimation for plant breeding using deep learning. *Computers and Electronics in Agriculture*, *150*, 196-204. DOI: https://doi.org/10.1016/j.compag.2018.04.024

Wei, M. C. F., & Molin, J. P. (2020). Soybean yield estimation and its components : A linear regression approach. *Agriculture*, *10*(8), 1-13. DOI: https://doi.org/10.3390/agriculture10080348

Xu, B., Wang, W., Falzon, G., Kwan, P., Guo, L., Chen, G., ... Schneider, D. (2020). Automated cattle counting using Mask R-CNN in quadcopter vision system. *Computers and Electronics in Agriculture*, *171*. DOI: https://doi.org/10.1016/j.compag.2020.105300