
Photo by P. Espitia/CIAT

AI-powered high-throughput phenotyping of seed yield in *Urochloa*: a deep learning approach

Darwin A. Arrechea-Castillo[1], Paula Espitia-Buitrago[1], Ronald D. Arboleda[1], David Florian-Vargas[2], Rosa N. Jauregui[1], Juan A. Cardoso[1].

[1]*International Center for Tropical Agriculture, Tropical Forages Program.* [2] *Grupo Papalotla, Research and Development Area.*

Contact: d.arrechea@cgiar.org

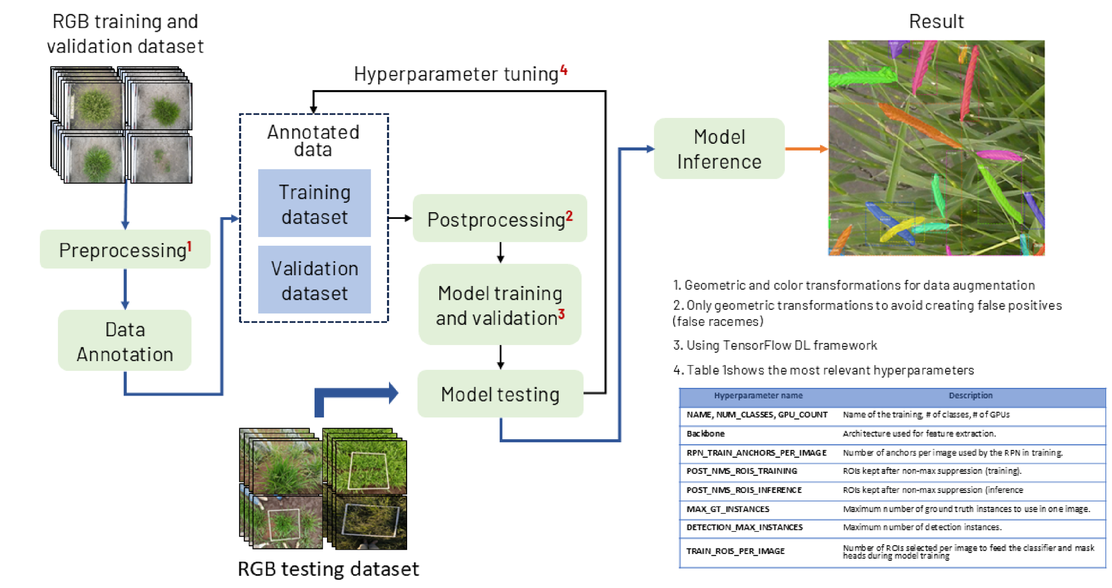
INTRODUCTION

- Measuring seed production components in a breeding scheme allows for the selection of genotypes with high seed yield potential[a].
- While the number of racemes is a key trait for estimating potential seed production, manual measurement across large populations is labor-intensive.
- **Challenges:** raceme overlap, large variations in raceme numbers per image, small raceme size relative to the overall image, and spectral signature similarities between the racemes and the rest of the plant.

OBJECTIVE

Develop an AI-based high-throughput phenotyping tool to accurately detect and count racemes in *Urochloa* spp. plots for its application in the breeding program.

METHODOLOGY



Hyperparameter name	Description
NAME, NUM_CLASSES, GPU_COUNT	Name of the training, # of classes, # of GPUs
Backbone	Architecture used for feature extraction.
RPN_TRAIN_ANCHORS_PER_IMAGE	Number of anchors per image used by the RPN in training.
POST_NMS_ROIS_TRAINING	ROIs kept after non-max suppression (training).
POST_NMS_ROIS_INFERENCE	ROIs kept after non-max suppression (inference
MAX_GT_INSTANCES	Maximum number of ground truth instances to use in one image.
DETECTION_MAX_INSTANCES	Maximum number of detection instances.
TRAIN_ROIS_PER_IMAGE	Number of ROIs selected per image to feed the classifier and mask heads during model training

Figure 1. Raceme detection and counting approach.

Training set: 2,400 nadir images were captured on 200 Urochloa genotypes, spaced 1.5 meters apart, over 7-months, to form a training set using a professional camera mounted on a buggy tractor (height of 1.2 m). The buggy was equipped with a light diffuser to ensure consistent, soft lighting a (Fig. 2).

- **Annotation process for training:** CVAT, a free web-based platform (Fig. 3).

Figure 2. Training and validation.

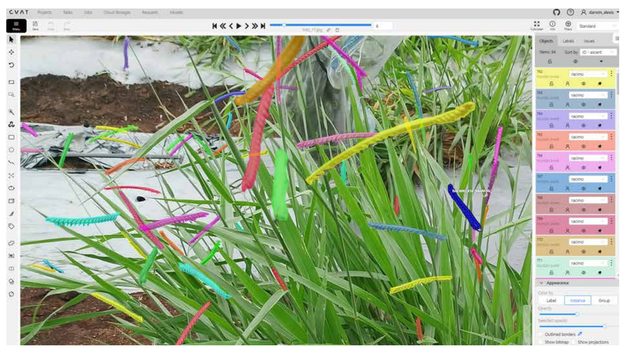


Figure 3. Annotation proccess.

Validation set:

- **Test Dataset 1:** 31 frontal view images were captured with a black background using Oukitel WP21 smartphone positioned at a fixed distance of 1.90 m at the CIAT campus in Palmira, Colombia (Fig. 4).
- **Test Dataset 2:** 24 images captured (Realme C53 smartphone) at varying distances from 1 to 2 m at the Tropical Forage Research Center (CIPAT) in Ocozocoautla de Espinosa, Chiapas, Mexico (Fig. 5).

Figure 4. Test Dataset 1.

Figure 5. Test Dataset 2.

RESULTS

Photointerpretation results - Test Dataset 1: All the images were annotated and processed by the model to compare its robustness against human instance segmentation, obtaining a precision of 96.16%, a recall of 94.10%, an F1-score of 95.12%, and a mean absolute error of 1.74 (Fig. 6).

In situ results - Testing 2 dataset: The model under-represents the total number of racemes in the images due to its 2D perspective, while human evaluators have a 3D view of the reality. However, the model categorized accurately plants based on the number of racemes (r = 0.88) (Fig. 7).

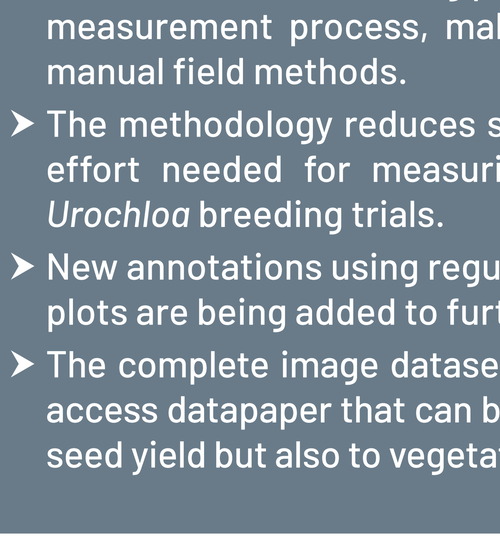


Figure 6. Raceme instance segmentation.

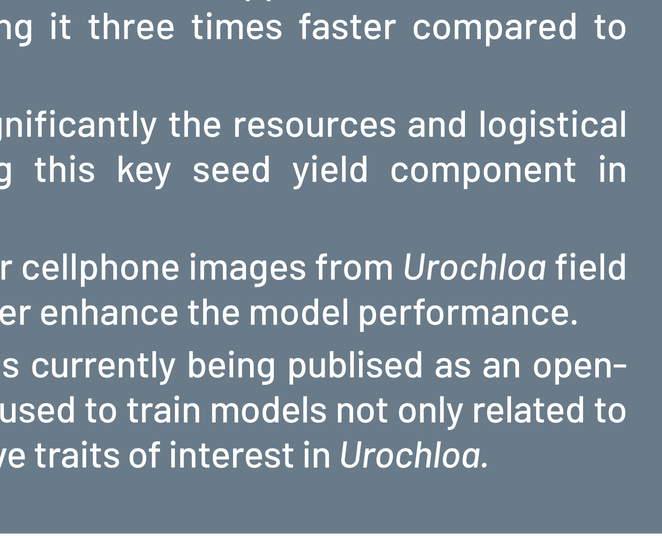


Figure 7. Raceme count comparison.

CONCLUSIONS

- The MaskRCNN model trained with 362 images with 25 567 annotations demonstrated high accuracy and precision for the raceme count in breeding plots of *Urochloa* spp. and accelerated the measurement process, making it three times faster compared to manual field methods.
- The methodology reduces significantly the resources and logistical effort needed for measuring this key seed yield component in *Urochloa* breeding trials.
- New annotations using regular cellphone images from *Urochloa* field plots are being added to further enhance the model performance.
- The complete image dataset is currently being publised as an open-access datapaper that can be used to train models not only related to seed yield but also to vegetative traits of interest in *Urochloa.*

REFERENCES

[a.] D. A. Arrechea-Castillo, P. Espitia-Buitrago, R. D. Arboleda, L. M. Hernandez, R. N. Jauregui, and J. A. Cardoso, "High-resolution image dataset for the automatic classification of phenological stage and identification of racemes in Urochloa spp. hybrids," Data Brief, article no. 110928, Sep. 2024, doi: 10.1016/j.dib.2024.110928.

ACKNOWLEDGEMENTS

This study was funded by the CGIAR Initiatives on Sustainable Animal Productivity, Livestock and Climate (L&C) and Accelerated Breeding. We also acknowledge the support of Grupo Papalotla. We thank all donors that globally support the work of the Initiatives through their contributions to the CGIAR system.

Alianza Bioversity & CIAT · CGIAR Initiative on Sustainable Animal Productivity · CGIAR Initiative on Livestock and Climate · CGIAR Initiative on Accelerated Breeding · Papalotla Semillas · IPPN · University of Nebraska Lincoln

GREEN HORIZONS — Navigating the future of plant phenotyping — IPPS Nº8 — Oct. 07.–11. 2024 Lincoln | Nebraska | USA

POSTER PREPARED FOR
8th International Plant Phenotyping Symposium 2024
October 7–11, 2024
Nebraska (USA)