# “End-to-end development of a mobile application to count the number of grains within a sorghum panicle using on-farm imagery”

by

Gustavo Nocera Santiago

B.S., Universidade de Sao Paulo, 2022

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Agronomy
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2024

Approved by

Major Professor
Ignacio A. Ciampitti

# Copyright

# Abstract

Estimating yield before harvesting can lead to important decisions related to crop management and use inputs. For the sorghum (*Sorghum bicolor L.*), yield estimation before harvest is laborious and time-consuming, mainly due to the fact that it relies on the panicle (head) being removed, threshed, and the grains being counted manually or machine-assisted. In this scenario, the integration of computer vision, artificial intelligence algorithms, and mobile applications represents a potential alternative to manual counting of grains on an on-farm scale. Consequently, this study aimed to develop an improved method for grain counting and, subsequently, yield estimation prior to harvesting, by combining artificial intelligence techniques with the development of a mobile application. Chapter 2 assesses the development of artificial intelligence techniques for the automated counting of grains in on-farm images captured using a mobile phone. In the aforementioned chapter, a total of 648 sorghum panicle images were benchmarked using a mobile phone in a field setting. Subsequently, the number of grains in 198 of these images was manually counted. Two frameworks were used to train segmentation models from the images: Yolov8 and Detectron2, with an average precision of 89% and 75%, respectively. Later, we tested three density models for the counting task: MCNN, TCNN-Seed, and Sorghum-Net (developed by this study), with the last overperforming the other two models with 17% of Mean Average Precision Error (MAPE). Lastly, a polynomial linear model was used to relate the counting obtained from the density model to the whole panicle, resulting in a MAPE of 17%. In Chapter 3, the best models from Chapter 2 (Yolov8, Sorghum-Net, and the polynomial model) were embedded into a mobile application for utilization in on-farm applications. We developed two branches of the mobile application: one operating locally and the other assisted by a web server. The user can obtain the yield estimation by inputting information on plant population and crop condition. A laboratory and a field validation were performed, yielding a relative root mean square error of 22.4% and 21.4%, respectively. The outcomes from this study can be employed by farmers, crop insurance companies, agronomists, breeders, and researchers to rapidly estimate and forecast sorghum grain yield before harvesting.

# Table of Contents

# List of Figures

# List of Tables

# Acknowledgments

I thank our Celestial Father for everything, my family Edson, Virginia, and Isabela for being my base and providing support to be here, my advisor Ignacio Ciampitti for the numerous opportunities, my girlfriend Giovanna, all the friends I made in Kansas for the emotional support, and Molly E. Brown for making an initial bridge between me and Ignacio, allowing my presence here. I also thank the Ciampitti Lab team for their assistance during data collection, mainly Ana Carcedo, for all the patience and support she gave me.

# Chapter 1 – Introduction

The sorghum crop (*Sorghum bicolor L.*) is a globally important cereal crop, ranking fifth in world production (C et al., 2017). Its economic importance derives from its versatility, which allows it to be used as human food, animal and poultry feed and as a biofuel (Hariprasanna & Rakshit, 2016). In addition to its numerous uses, sorghum has multiple advantages, including adaptability to drought, extreme temperatures, reduced input costs, and the potential to mitigate climate change effects (Srinivasa Rao et al., 2014). These factors contribute to the expectation that the future demand for sorghum will increase due to climate change, diversification in food alternatives in developing countries, and more diversified industrial uses (Srinivasa Rao et al., 2014).

The estimation of yields prior to the harvesting of crops is of paramount importance for farmers, as it allows them to make informed decisions regarding the vegetative development of the crop and the management of resources (Raj et al., 2023). For grain sorghum, the estimation of yield is mainly driven by three key factors i) plant density (number of plants per unit area), ii) number of grains per panicle, and iii) grain weight (Ciampitti and Carcedo, 2023). Of the three, the most challenging is obtaining a reliable estimate on the number of grains per panicle, which relies on laborious, time-consuming, and destructive steps when this process is done via manual counting.

The application of computer vision and deep learning techniques has been explored in the literature as a means of assisting with the counting of agricultural products, with notable success. Recent reviews have highlighted the use of deep learning for object counting in agriculture, presenting better metrics than traditional computer vision techniques (Farjon et al., 2023; Santos et al., 2020). These deep-learning models have been largely applied for fruit counting and yield estimations (Albahar, 2023; Gupta, 2021). In the context of grain sorghum crops, the integration of deep-learning techniques with unmanned aerial vehicle (UAV) imagery has facilitated the enumeration of panicle populations in agricultural fields and the estimation of their weight (Zannou & Houndji, 2019). However, this approach has not been extended to assess the number of grains within sorghum panicles .

Mobile applications represent an affordable and non-destructive methodology for estimating yield through the utilization of deep learning models. Examples of mobile applications developed to facilitate agricultural practices and estimate yield can be found in the current scientific literature (Mendes et al., 2020). Deep learning techniques and semantic segmentation have been applied to address issues such as fruit detection and counting in orchards, overcoming challenges such as occlusion and illumination variations (Maheswari et al., 2021). In the context of grain sorghum, previous studies (Ciampitti et al., 2015) have conducted a study to develop a mobile application for estimation of grain yield. This involved the positioning of the sorghum panicle, with a standard square of known dimensions, against a background of known volume, with the objective of estimating the harvestable organ volume. Nevertheless, it is not currently possible to use a mobile application to count the grains of a sorghum panicle using on-farm image data in order to predict sorghum yield prior to the harvest.

This thesis has a structure divided into chapters. Chapter 2 presents the development of the deep learning and linear model used to segment, count, and estimate the number of grains in an image. Chapter 3 presents the mobile application developed to employ the models in on-farm image scenarios. The specific goals for Chapter 2 were to i) detect and segment sorghum panicles at field scale using imagery from a smartphone camera, ii) utilize segmented imagery to count the number of grains within manually segmented panicles, and iii) employ a linear model to estimate the total number of grains within a panicle based on the previous counting. The objectives of Chapter 3 were to develop a mobile application that counts the number of grains within a sorghum panicle using field scale imagery, herein termed sorghum panicle photometry. The application demonstrates the potential to assist in forecasting yields through the integration of user inputs and the output obtained from sorghum images (grain numbers), ultimately generating a report for the user with an overall sorghum yield estimation.

# References

Abstract: Development of a New Tool for Estimating Sorghum Yields at the Farm-Scale. (ASA, CSSA and SSSA International Annual Meetings). (n.d.). Retrieved July 19, 2023, from https://scisoc.confex.com/scisoc/2014am/webprogram/Paper88108.html

Abstract: Sorghumyield® App, a New Tool for Predicting Sorghum Yields: Validation and Calibration. (ASA, CSSA and SSSA International Annual Meetings (2015)). (n.d.). Retrieved March 6, 2024, from https://scisoc.confex.com/crops/2015am/webprogram/Paper93787.html

Albahar, M. (2023). A Survey on Deep Learning and Its Impact on Agriculture: Challenges and Opportunities. Agriculture, 13(3), Article 3. https://doi.org/10.3390/agriculture13030540

Araus, J. L., Kefauver, S. C., Zaman-Allah, M., Olsen, M. S., & Cairns, J. E. (2018). Translating High-Throughput Phenotyping into Genetic Gain. Trends in Plant Science, 23(5), 451–466. https://doi.org/10.1016/j.tplants.2018.02.001

Arya, S., Sandhu, K. S., Singh, J., & kumar, S. (2022). Deep learning: As the new frontier in high-throughput plant phenotyping. Euphytica, 218(4), 47. https://doi.org/10.1007/s10681-022-02992-3

Boominathan, L., Kruthiventi, S. S. S., & Babu, R. V. (2016). CrowdNet: A Deep Convolutional Network for Dense Crowd Counting. Proceedings of the 24th ACM International Conference on Multimedia, 640–644. https://doi.org/10.1145/2964284.2967300

Bradski, G. (2000). The OpenCV Library. Dr. Dobb’s Journal: Software Tools for the Professional Programmer, 25(11), 120–123.

C, A., Rao, B. D., & Rao, V. A. T. and T. G. N. (2017). Improving production and utilization of sorghum in Asia Indian Institute of Millets Research, India. In Achieving sustainable cultivation of sorghum Volume 2. Burleigh Dodds Science Publishing.

Cai, Z., & Vasconcelos, N. (2017). Cascade R-CNN: Delving into High Quality Object Detection (arXiv:1712.00726). arXiv. https://doi.org/10.48550/arXiv.1712.00726

Cao, Z., Jiang, K., Zhou, W., Xu, S., Peng, H., & Yang, D. (2023). Continuous improvement of self-driving cars using dynamic confidence-aware reinforcement learning. Nature Machine Intelligence, 5(2), 145–158. https://doi.org/10.1038/s42256-023-00610-y

Cereals & Grains | Inpho | Food and Agriculture Organization of the United Nations. (n.d.). Retrieved July 19, 2023, from https://www.fao.org/in-action/inpho/crop-compendium/cereals-grains/en/

Davies, E. R. (2005). CHAPTER 1—Vision, the Challenge. In E. R. Davies (Ed.), Machine Vision (Third Edition) (pp. 1–15). Morgan Kaufmann. https://doi.org/10.1016/B978-012206093-9/50004-6

Demarco, P. A., Mayor, L., Rotundo, J. L., Prasad, P. V. V., Morris, G. P., Fernandez, J. A., Tamagno, S., Hammer, G., Messina, C. D., & Ciampitti, I. A. (2023). Retrospective study in U.S. commercial sorghum breeding: II. Physiological changes associated to yield gain. Crop Science, 63(2), 867–878. https://doi.org/10.1002/csc2.20845

Deng, J., Dong, W., Socher, R., Li, L.-J., Kai Li, & Li Fei-Fei. (2009). ImageNet: A large-scale hierarchical image database. 2009 IEEE Conference on Computer Vision and Pattern Recognition, 248–255. https://doi.org/10.1109/CVPR.2009.5206848

Facebookresearch/detectron2. (2023). [Python]. Meta Research. https://github.com/facebookresearch/detectron2 (Original work published 2019)

Farjon, G., Huijun, L., & Edan, Y. (2023). Deep-learning-based counting methods, datasets, and applications in agriculture: A review. Precision Agriculture, 24(5), 1683–1711. https://doi.org/10.1007/s11119-023-10034-8

Fountas, S., Carli, G., Sørensen, C. G., Tsiropoulos, Z., Cavalaris, C., Vatsanidou, A., Liakos, B., Canavari, M., Wiebensohn, J., & Tisserye, B. (2015). Farm management information systems: Current situation and future perspectives. Computers and Electronics in Agriculture, 115, 40–50. https://doi.org/10.1016/j.compag.2015.05.011

Ghafari, M., Mailman, D., Hatami, P., Peyton, T., Yang, L., Dang, W., & Qin, H. (2022). A Comparison of YOLO and Mask-RCNN for Detecting Cells from Microfluidic Images. 2022 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), 204–209. https://doi.org/10.1109/ICAIIC54071.2022.9722616

Ghosal, S., Zheng, B., Chapman, S. C., Potgieter, A. B., Jordan, D. R., Wang, X., Singh, A. K., Singh, A., Hirafuji, M., Ninomiya, S., Ganapathysubramanian, B., Sarkar, S., & Guo, W. (2019). A Weakly Supervised Deep Learning Framework for Sorghum Head Detection and Counting. Plant Phenomics, 2019. https://doi.org/10.34133/2019/1525874

Güler, R. A., Neverova, N., & Kokkinos, I. (2018). DensePose: Dense Human Pose Estimation In The Wild (arXiv:1802.00434). arXiv. https://doi.org/10.48550/arXiv.1802.00434

Gupta, P. (2021). APPLICATIONS OF DEEP LEARNING METHODS IN AGRICULTURE: A REVIEW. 8.

Haider, T., Farid, M. S., Mahmood, R., Ilyas, A., Khan, M. H., Haider, S. T.-A., Chaudhry, M. H., & Gul, M. (2021). A Computer-Vision-Based Approach for Nitrogen Content Estimation in Plant Leaves. Agriculture, 11(8), Article 8. https://doi.org/10.3390/agriculture11080766

Hariprasanna, K., & Rakshit, S. (2016). Economic Importance of Sorghum. In S. Rakshit & Y.-H. Wang (Eds.), The Sorghum Genome (pp. 1–25). Springer International Publishing. https://doi.org/10.1007/978-3-319-47789-3_1

Harris, C. R., Millman, K. J., Van Der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., Van Kerkwijk, M. H., Brett, M., Haldane, A., Del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. Nature, 585(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2018). Mask R-CNN (arXiv:1703.06870). arXiv. https://doi.org/10.48550/arXiv.1703.06870

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition (arXiv:1512.03385). arXiv. https://doi.org/10.48550/arXiv.1512.03385

Hu, H., Kaizu, Y., Zhang, H., Xu, Y., Imou, K., Li, M., Huang, J., & Dai, S. (2022). Recognition and localization of strawberries from 3D binocular cameras for a strawberry picking robot using coupled YOLO/Mask R-CNN. International Journal of Agricultural and Biological Engineering, 15(6), Article 6. https://doi.org/10.25165/ijabe.v15i6.7306

Hu, H., Tang, C., Shi, C., & Qian, Y. (2023). Detection of residual feed in aquaculture using YOLO and Mask RCNN. Aquacultural Engineering, 100, 102304. https://doi.org/10.1016/j.aquaeng.2022.102304

James, C., Smith, D., He, W., Chandra, S. S., & Chapman, S. C. (2024). GrainPointNet: A deep-learning framework for non-invasive sorghum panicle grain count phenotyping. Computers and Electronics in Agriculture, 217, 108485. https://doi.org/10.1016/j.compag.2023.108485

Jocher, G., Chaurasia, A., & Qiu, J. (2023). YOLO by Ultralytics (Version 8.0.0) [Python]. https://github.com/ultralytics/ultralytics (Original work published 2022)

Jung, A. (2023). Imgaug [Python]. https://github.com/aleju/imgaug (Original work published 2015)

Kalischuk, M., Paret, M. L., Freeman, J. H., Raj, D., Da Silva, S., Eubanks, S., Wiggins, D. J., Lollar, M., Marois, J. J., Mellinger, H. C., & Das, J. (2019). An Improved Crop Scouting Technique Incorporating Unmanned Aerial Vehicle–Assisted Multispectral Crop Imaging into Conventional Scouting Practice for Gummy Stem Blight in Watermelon. Plant Disease, 103(7), 1642–1650. https://doi.org/10.1094/PDIS-08-18-1373-RE

Khaki, S., Pham, H., Han, Y., Kuhl, A., Kent, W., & Wang, L. (2021). DeepCorn: A semi-supervised deep learning method for high-throughput image-based corn kernel counting and yield estimation. Knowledge-Based Systems, 218, 106874. https://doi.org/10.1016/j.knosys.2021.106874

Kristiani, E., Yang, C.-T., & Phuong Nguyen, K. L. (2020). Optimization of Deep Learning Inference on Edge Devices. 2020 International Conference on Pervasive Artificial Intelligence (ICPAI), 264–267. https://doi.org/10.1109/ICPAI51961.2020.00056

Lane, N. D., Bhattacharya, S., Georgiev, P., Forlivesi, C., Jiao, L., Qendro, L., & Kawsar, F. (2016). DeepX: A Software Accelerator for Low-Power Deep Learning Inference on Mobile Devices. 2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN), 1–12. https://doi.org/10.1109/IPSN.2016.7460664

Li, H., Chen, Z., Liu, G., Jiang, Z., & Huang, C. (2017). Improving Winter Wheat Yield Estimation from the CERES-Wheat Model to Assimilate Leaf Area Index with Different

Assimilation Methods and Spatio-Temporal Scales. Remote Sensing, 9(3), Article 3. https://doi.org/10.3390/rs9030190

Li, H., Wang, P., & Huang, C. (2022). Comparison of Deep Learning Methods for Detecting and Counting Sorghum Heads in UAV Imagery. Remote Sensing, 14(13), Article 13. https://doi.org/10.3390/rs14133143

Li, M., Shao, M.-R., Zeng, D., Ju, T., Kellogg, E. A., & Topp, C. N. (2020). Comprehensive 3D phenotyping reveals continuous morphological variation across genetically diverse sorghum inflorescences. New Phytologist, 226(6), 1873–1885. https://doi.org/10.1111/nph.16533

Li, Y., Jia, J., Zhang, L., Khattak, A. M., Sun, S., Gao, W., & Wang, M. (2019). Soybean Seed Counting Based on Pod Image Using Two-Column Convolution Neural Network. IEEE Access, 7, 64177–64185. https://doi.org/10.1109/ACCESS.2019.2916931

Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2018). Focal Loss for Dense Object Detection (arXiv:1708.02002). arXiv. https://doi.org/10.48550/arXiv.1708.02002

Lin, T.-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., & Dollár, P. (2015). Microsoft COCO: Common Objects in Context (arXiv:1405.0312). arXiv. https://doi.org/10.48550/arXiv.1405.0312

Lin, Z., & Guo, W. (2020). Sorghum Panicle Detection and Counting Using Unmanned Aerial System Images and Deep Learning. Frontiers in Plant Science, 11. https://www.frontiersin.org/articles/10.3389/fpls.2020.534853

Lipper, L., Thornton, P., Campbell, B. M., Baedeker, T., Braimoh, A., Bwalya, M., Caron, P., Cattaneo, A., Garrity, D., Henry, K., Hottle, R., Jackson, L., Jarvis, A., Kossam, F., Mann, W., McCarthy, N., Meybeck, A., Neufeldt, H., Remington, T., … Torquebiau, E.

F. (2014). Climate-smart agriculture for food security. Nature Climate Change, 4(12), Article 12. https://doi.org/10.1038/nclimate2437

Maheswari, P., Raja, P., Apolo-Apolo, O. E., & Pérez-Ruiz, M. (2021). Intelligent Fruit Yield Estimation for Orchards Using Deep Learning Based Semantic Segmentation Techniques—A Review. Frontiers in Plant Science, 12. https://doi.org/10.3389/fpls.2021.684328

Malambo, L., Popescu, S., Ku, N.-W., Rooney, W., Zhou, T., & Moore, S. (2019). A Deep Learning Semantic Segmentation-Based Approach for Field-Level Sorghum Panicle Counting. Remote Sensing, 11(24), Article 24. https://doi.org/10.3390/rs11242939

Mendes, J., Pinho, T. M., Neves dos Santos, F., Sousa, J. J., Peres, E., Boaventura-Cunha, J., Cunha, M., & Morais, R. (2020). Smartphone Applications Targeting Precision Agriculture Practices—A Systematic Review. Agronomy, 10(6), Article 6. https://doi.org/10.3390/agronomy10060855

Monk, R. L., Miller, F. R., & McBee, G. G. (1984). Sorghum improvement for energy production. Biomass, 6(1), 145–153. https://doi.org/10.1016/0144-4565(84)90017-9

Oh, M., Olsen, P., & Ramamurthy, K. N. (2019). Counting and Segmenting Sorghum Heads (arXiv:1905.13291). arXiv. https://doi.org/10.48550/arXiv.1905.13291

Patel, A. K., Park, E.-S., Lee, H., Priya, G. G. L., Kim, H., Joshi, R., Arief, M. A. A., Kim, M. S., Baek, I., & Cho, B.-K. (2023). Deep Learning-Based Plant Organ Segmentation and Phenotyping of Sorghum Plants Using LiDAR Point Cloud. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 16, 8492–8507. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. https://doi.org/10.1109/JSTARS.2023.3312815

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., & Cournapeau, D. (n.d.). Scikit-learn: Machine Learning in Python. MACHINE LEARNING IN PYTHON.

Petrellis, N. (2017). A smart phone image processing application for plant disease diagnosis. 2017 6th International Conference on Modern Circuits and Systems Technologies (MOCAST), 1–4. https://doi.org/10.1109/MOCAST.2017.7937683

Prasetyo, E., Suciati, N., & Fatichah, C. (2020). A Comparison of YOLO and Mask R-CNN for Segmenting Head and Tail of Fish. 2020 4th International Conference on Informatics and Computational Sciences (ICICoS), 1–6. https://doi.org/10.1109/ICICoS51170.2020.9299024

Raj, H. C., Boomiraj, K., Pazhanivelan, S., Moorthy, N. K. S., & Prasanthrajan, M. (2023). Spatial Yield Assessment of Sorghum (Sorghum Bicolor) Using DSSAT – CERES – Sorghum Model in Solapur, Maharashtra, India. International Journal of Environment and Climate Change, 13(10), 546–554. https://doi.org/10.9734/ijecc/2023/v13i102684

Robinson, J., Jegelka, S., & Sra, S. (2020). Strength from Weakness: Fast Learning Using Weak Supervision. Proceedings of the 37th International Conference on Machine Learning, 8127–8136. https://proceedings.mlr.press/v119/robinson20a.html

Roboflow: Give your software the power to see objects in images and video. (n.d.). Retrieved July 25, 2023, from https://roboflow.com/

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation (arXiv:1505.04597). arXiv. https://doi.org/10.48550/arXiv.1505.04597

Rühling Cachay, S., Boecking, B., & Dubrawski, A. (2021). End-to-End Weak Supervision. Advances in Neural Information Processing Systems, 34, 1845–1857. https://proceedings.neurips.cc/paper_files/paper/2021/hash/0e674a918ebca3f78bfe02e2f387689d-Abstract.html

Santiago, G., Carcedo, A., Marziotte, L., & Ciampitti, I. (2023). Algorithm to Estimate Sorghum Grain Number from Panicles Using Images Collected with a Smartphone at Field-Scale. Kansas Agricultural Experiment Station Research Reports, 9(4). https://doi.org/10.4148/2378-5977.8465

Santos, L., Santos, F. N., Oliveira, P. M., & Shinde, P. (2020). Deep Learning Applications in Agriculture: A Short Review. In M. F. Silva, J. Luís Lima, L. P. Reis, A. Sanfeliu, & D. Tardioli (Eds.), Robot 2019: Fourth Iberian Robotics Conference (pp. 139–151). Springer International Publishing. https://doi.org/10.1007/978-3-030-35990-4_12

Srinivasa Rao, P., Reddy, B. V. S., Nagaraj, N., & Upadhyaya, H. D. (2014). Sorghum Production for Diversified Uses (Y. H. Wang, H. D. Upadhyaya, & C. Kole, Eds.; pp. 01–27). CRC Press (Taylor & Francis). http://www.crcnetbase.com/doi/abs/10.1201/b17153-2

van Oosterom, E. J., & Hammer, G. L. (2008). Determination of grain number in sorghum. Field Crops Research, 108(3), 259–268. https://doi.org/10.1016/j.fcr.2008.06.001

Velasco-Montero, D., Femández-Bemi, J., Carmona-Gálán, R., & Rodríguez-Vázquez, A. (2019). On the Correlation of CNN Performance and Hardware Metrics for Visual Inference on a Low-Cost CPU-based Platform. 2019 International Conference on Systems, Signals and Image Processing (IWSSIP), 249–254. https://doi.org/10.1109/IWSSIP.2019.8787329

Véstias, M. P. (2019). A Survey of Convolutional Neural Networks on Edge with Reconfigurable Computing. Algorithms, 12(8), Article 8. https://doi.org/10.3390/a12080154

Vieyra, R. E., Megowan-Romanowicz, C., O'Brien, D. J., Vieyra, C., & Johnson-Glenberg, M. C. (2023). Harnessing the Digital Science Education Revolution: Smartphone Sensors as Teaching Tools. In Theoretical and Practical Teaching Strategies for K-12 Science Education in the Digital Age (pp. 131–152). IGI Global. https://doi.org/10.4018/978-1-6684-5585-2.ch008

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., Van Der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … Vázquez-Baeza, Y. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. Nature Methods, 17(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2

Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep Learning for Computer Vision: A Brief Review. Computational Intelligence and Neuroscience, 2018, e7068349. https://doi.org/10.1155/2018/7068349

Wan, S., & Goudos, S. (2020). Faster R-CNN for multi-class fruit detection using a robotic vision system. Computer Networks, 168, 107036. https://doi.org/10.1016/j.comnet.2019.107036

Wang, J., Lin, J., & Wang, Z. (2018). Efficient Hardware Architectures for Deep Convolutional Neural Network. IEEE Transactions on Circuits and Systems I: Regular Papers, 65(6), 1941–1953. IEEE Transactions on Circuits and Systems I: Regular Papers. https://doi.org/10.1109/TCSI.2017.2767204

Wiley, V., & Lucas, T. (2018). Computer Vision and Image Processing: A Paper Review. International Journal of Artificial Intelligence Research, 2(1), Article 1. https://doi.org/10.29099/ijair.v2i1.42

Yang, C., Everitt, J. H., & Bradford, J. M. (2009). Evaluating high resolution SPOT 5 satellite imagery to estimate crop yield. Precision Agriculture, 10(4), 292–303. https://doi.org/10.1007/s11119-009-9120-6

Ye, X., Abe, S., Zhang, S., & Yoshimura, H. (2020). Rapid and non-destructive assessment of nutritional status in apple trees using a new smartphone-based wireless crop scanner system. Computers and Electronics in Agriculture, 173, 105417. https://doi.org/10.1016/j.compag.2020.105417

Young, S. N., Kayacan, E., & Peschel, J. M. (2019). Design and field evaluation of a ground robot for high-throughput phenotyping of energy sorghum. Precision Agriculture, 20(4), 697–722. https://doi.org/10.1007/s11119-018-9601-6

Zannou, J. G. N., & Houndji, V. R. (2019). Sorghum Yield Prediction using Machine Learning. 2019 3rd International Conference on Bio-Engineering for Smart Technologies (BioSMART), 1–4. https://doi.org/10.1109/BIOSMART.2019.8734219

Zhang, Y., Zhou, D., Chen, S., Gao, S., & Ma, Y. (2016). Single-Image Crowd Counting via Multi-Column Convolutional Neural Network. 589–597. https://openaccess.thecvf.com/content_cvpr_2016/html/Zhang_Single-Image_Crowd_Counting_CVPR_2016_paper.html

# Chapter 2 – Deep learning methods using imagery from a smartphone for recognizing sorghum panicles and counting grains at a plant level

*Gustavo N. Santiago, Pedro H. Cisdeli Magalhaes, Ana J. P. Carcedo, Lucia Marziotte, Laura Mayor, Ignacio A. Ciampitti. Deep Learning Methods Using Imagery from a Smartphone for Recognizing Sorghum Panicles and Counting Grains at a Plant Level. Plant Phenomics. 2024;6:0234.DOI:10.34133/plantphenomics.0234*

## Abstract

High-throughput phenotyping is the bottleneck for advancing field trait characterization and yield improvement in major field crops. Specifically for sorghum (*Sorghum bicolor* L.), rapid plant-level yield estimation is highly dependent on characterizing the number of grains within a panicle. In this context, the integration of computer vision and artificial intelligence algorithms with traditional field phenotyping can be a critical solution to reduce labor costs and time. Therefore, this study aims to improve sorghum panicle detection and of grain number estimation from smartphone-capture images under field conditions. A pre-harvest benchmark dataset was collected at field-scale (2023 season, KS, US), with 648 images of sorghum panicles retrieved via smartphone device, and grain number counted. Each sorghum panicle image was manually labeled, and the images were augmented. Two models were trained using the Detectron2 and Yolov8 frameworks for detection and segmentation, with an average precision of 75% and 89%, respectively. For the grain number, three models were trained: MCNN, TCNN-Seed, and Sorghum-Net (developed in this study). The Sorghum-Net model showed a mean absolute percentage error of 17%, surpassing the other models. Lastly, a simple equation was presented to relate the count from the model (using images from only one side of the panicle) to the field-derived observed number of grains per sorghum panicle. The resulting framework

obtained an estimation of grain number with a 17% error. The proposed framework lays the foundation for the development of a more robust application to estimate sorghum yield using images from a smartphone at the plant level.

# Introduction

Sorghum (*Sorghum bicolor* L.) plays a pivotal role in global food security, serving as a vital source of nutrition for both humans and livestock in arid regions (FAO, 2023; Monk et al., 1984). Accurate production estimates are crucial to informed decision making by farmers, guiding choices on crop and financial management decisions (Fountas et al., 2015). In addition, yield predictions are essential for breeders when selecting new genotypes. However, the conventional method of sorghum yield estimation relies on labor-intensive and time-consuming manual sampling at the plant level, which introduces the potential for human error, especially in large areas or with a large number of plots (Z. Lin & Guo, 2020).

To address these challenges, there is a pressing need to develop high-throughput phenotyping methods to automate sorghum yield estimation. Grain number is a key component of sorghum yield (van Oosterom & Hammer, 2008), with a few attempts to estimate grain numbers using traditional and image-based approaches, but have not been effectively translated to field scale (Ciampitti et al., 2014; Ciampitti et al., 2015; Santiago et al., 2023). A promising alternative is to use computer vision algorithms to extract meaningful information from images and videos (Davies, 2005; Wan & Goudos, 2020). Computer vision encompasses computational techniques for retrieving desired data through image processing (Wiley & Lucas, 2018), with convolutional neural networks (CNNs) being a notable image processing method inspired by the human neural system (Voulodimos et al., 2018).

More recently, the application of CNN techniques in agriculture, especially for high-throughput phenotyping, has gained traction. Arya et al. (Arya et al., 2022) explored how deep learning and CNN techniques can help in plant phenotyping to speed up the current plant trait identification and selection process. James et al. (2024), M. Li et al. (2020), Patel et al. (2023) developed 3D, X-ray computed tomography, point cloud, and CNN models that are relevant for analyzing phenotypic aspects of sorghum plants, such as panicle architecture, grain number, and organ segmentation. However, a major challenge for many of these methods is the effective and pragmatic translation of these efforts to field conditions. To bring these techniques to the field scale, Young et al. (2019) developed a ground robot capable of performing high-throughput phenotyping of sorghum plants. However, the relatively high cost of this technology remains a limitation to scalability and adoption.

With the aim of estimating sorghum yield production, several studies have utilized CNN models, such as two-stage CNN (Oh et al., 2019), U-Net (Ronneberger et al., 2015), weakly supervised frameworks (Ghosal et al., 2019), RetinaNet (T.-Y. Lin et al., 2018), and ResNet-50 (He et al., 2015), to accurately count sorghum panicles from imagery, achieving impressive results, often exceeding 95% accuracy. However, these studies predominantly relied on imagery obtained through unmanned aerial vehicles (UAVs), which presents challenges such as high equipment costs and the need for specialized training. To overcome these limitations, the use of smartphones, with their increasing computing power, embedded sensors, and portability (Vieyra et al., 2023), offers a viable solution. The current trend of using smartphone cameras in agriculture has seen success in various applications, including estimating plant nitrogen content and plant health in different crops (Haider et al., 2021; Petrellis, 2017; Ye et al., 2020). However, efforts to estimate sorghum yield using smartphones under field conditions have been limited, with previous attempts using traditional image manipulation techniques (Ciampitti et al., 2015; Santiago et al., 2023).

This study aims to address this research knowledge gap by using CNNs to effectively remove background under plant-level conditions and count sorghum grains. Therefore, the specific objectives include i) detecting and segmenting sorghum panicles at the field level using images from a smartphone camera, ii) using segmented images to count the number of grains within manually segmented panicles, and iii) using a linear model to estimate the total number of grains within a panicle based on the previous count. Through this innovative approach, we aim to provide a practical and cost-effective solution to pave the way for the development of a rapid sorghum yield estimation at the plant level, advancing agronomy and computer science.

## Materials and Methods

The study has been conducted employing a Lenovo Legion desktop computer with an Intel® Core™ i7-10700 CPU @ 2.90GHz, NVIDIA GeForce RTX 2070 SUPER 8GB; 16GB RAM using CUDA 11.8 and Python 3.9. The overall description of the proposed framework is presented in Figure 2.1.

### Detection and Segmentation

*Data*A total of 457 images of sorghum panicles from 20 commercial hybrids were obtained during the pre-harvest (October 14, 2023; 128 days after sowing), at an experimental plot located at Wamego, KS, US from 9 to 12 a.m. The 20 commercial sorghum hybrids encompassed a wide variability in their canopy and panicle architecture, panicle colors, and maturity groups, including 18 genotypes proceeding from the US breeding program and two tropical hybrids from the Mexican breeding program. The images centered on a single panicle at a distance of one meter distance and parallel to the panicle, with other sorghum plants in the background. A generic smartphone was used with a camera of 64MP primary camera based on Sony IMX 582 Quad-Bayer 1/2" sensor with 0.8µm pixels, 25mm f/1.8 lens, set to automatic mode of image capture and saved in JPEG file extension. In all images, the sunlight was in the same direction as the camera to avoid high brightness in the camera sensor, and the distance to the panicles was 1 m, as shown in Figure 2.2.

Using Roboflow software (*Roboflow*, 2023), all panicles in the framed images were manually labeled with polygons. The only class considered in this study was "panicles", which refers only to the sorghum panicles, excluding the rest. An augmentation step was used to increase the image dataset. The images were subjected to random horizontal and vertical flipping, horizontal and vertical shifting, color and brightness modifications, zooming, and random cropping. Finally, the complete dataset resulted in 4,623 images of 640x640 pixels in size, with 4,500 for training, 102 for validation, and 21 for visual inspection. Some examples of the images from the resulting dataset are shown in Figure 2.3.

*Frameworks, parameters, and metrics*

Two frameworks were compared in this step: Detectron2 (*Facebookresearch/Detectron2*, 2019) and Yolov8 (Jocher et al., 2023). Detectron2 was developed by Meta AI Research as the successor of Detectron and maskrcnn-bechmark and has several pre-trained models for different computer vision applications. The Detectron2 framework includes models such as Faster R-CNN (Wan & Goudos, 2020), Mask R-CNN (He et al., 2018), RetinaNet (T.-Y. Lin et al., 2018), DensePose (Güler et al., 2018), and Cascade R-CNN (Cai & Vasconcelos, 2017). The Detectron2 CNN architecture used in this study was Mask R-CNN (He et al., 2018). Yolov8, a member of the You Only Look Once (YOLO) family, was developed by Ultralytics for detection, segmentation, classification, pose, and tracking tasks. It has models pre-trained on COCO (T.-Y. Lin et al., 2015), Open Image V7, and ImageNet (Deng et al., 2009) datasets. For further details on all model architectures can be found in the references presented in this section.

The following parameters were employed to train the models for the detection and segmentation in both frameworks: 28 epochs, a learning rate of 0.001, and a batch size of 8. For the other parameters, default values from the frameworks were employed. The pre-trained weights mask_rcnn_R_101_FPN_3x for Detectron2 and yolov8x-seg for Yolov8 were used to support the training process. The model outputs from the training were evaluated using the following metrics: average precision with 50% of the region compared to the ground truth (AP50) and average recall (AR), which were obtained directly from the frameworks. Finally, a visual inspection of 21 images from the dataset was performed to evaluate the overall performance of the segmentation task.

## Density maps

### *Data*

From the 457 images collected in the field, 40 images were randomly selected. The framed panicle within each image was manually segmented, excluding the background. Manual segmentation was used to ensure that the panicle of interest was perfectly segmented. The images were then cropped into 3 equal height areas: the top, middle, and bottom of the panicle. This step was necessary due to the large number of grains within a panicle (up to 4000 grains) and their small size (~8 pixels). This procedure resulted in 120 images.

Following the methodology proposed by Zhang et al. (Zhang et al., 2016), density maps were used to obtain the number of grains because generally, as detection methods in general have poor performance in detecting small objects (H. Li et al., 2022) and poor stability and scalability (Y. Li et al., 2019). Foreground segmentation is challenging, and an inaccurate segmentation could undermine the final count. The density maps provide the distribution of the number of grains in the image, allowing the researcher to corroborate the output performance.

The ground truth density maps were created by labeling the images by placing a point in each visible grain using the LabelMe software. This process generated a JSON file containing the point coordinates for each picture. The images and the point coordinates were then resized to 224x224. The density maps were created using the points coordinates and the function gaussian_filter from scipy (Virtanen et al., 2020) with a sigma value of five. The function sum from the numpy library (Harris et al., 2020) was employed to calculate the integral of all the pixels in the density map image and obtain the total amount of grains within an image. The images and the ground truth were augmented using imgaug (Jung, 2015) and opencv (Bradski, 2000) libraries by changing the hue and saturation, adding Gaussian noise and blur, inverting color channels, changing contrast, adding value to each pixel, rotating, and flipping the images. At the end of the augmentation process, the dataset was composed of 4800 images, which were split into 80% training and 20% testing.

*Model architectures, parameters, and metrics*

To train the models, three different architectures were employed: MCNN (Zhang et al., 2016), TCNN-Seed (Y. Li et al., 2019) and Sorghum-Net. The MCNN is a multi-column CNN developed to count the number of people in an image, obtaining a good performance in crowd images. The TCNN-Seed was developed to count the number of seeds based on pod images. TCNN-Seed is a two-column CNN that presents low error and a lower number of parameters, compared to MCNN. For more information regarding the architecture of these two CNN models, please refer to the original literature. Sorghum-Net is a multi-column CNN that was specifically developed for this study, composed of three columns with four convolutional layers each, a concatenation, an up-sampling, and a convolutional layer at the end. The first column has a kernel size of 3x3 with convolutional layer dimensions of 80, 160, 80, and 40. The second column has a kernel size of 5x5 with convolutional layer dimensions of 40, 80, 40, and 20. The third column has a kernel size of 7x7 with convolutional layer dimensions of 20, 40, 20, and 40. This model architecture uses batch normalization after each convolutional layer of the columns and in front of the activation function. The activation function is the rectified linear unit (ReLU). A max-pooling layer of a 2x2 size kernel with strides of 2 was used after the two first convolutional layers in each column.

To summarize, the columns presented similar structures as follows: Conv2d-MaxPooling-Conv2d-MaxPooling-Conv2d-Conv2d-Conv2d. After the parallel columns, the feature map was concatenated, and up-sampled to have the same size as the ground-truth density maps (224x224), and the up-sampled was used as input to a convolutional layer of 1x1 kernel size and a dimension of 2. The Sorghum-Net architecture is illustrated in Figure 2.4.

The TensorFlow framework was used to create and train the models. The parameters to train the models were set to a maximum epoch number of 200 with a patience of 20 epochs, batch size of 8, a learning rate of 0.001, the optimizer employed was 'Adam', and a loss function of Mean Squared Error (MSE).

The metrics used to compare the different models were: Mean Absolute Error (MAE) (equation 1), mean squared error (MSE) (equation 2), and Mean Absolute Percentage Error (MAPE) (equation 3), from the scipy library [50]. The values used in the metrics were obtained by comparing the number of grains predicted by integrating the pixels of the output density maps from the models and the number of grains observed (manually counted) in the respective imagery from the test dataset.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{1}$$

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{2}$$

$$\text{MAPE} = \left(\frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{y_i}\right) * 100 \tag{3}$$

Where n is the number of observations, $y_i$ is the actual value of the i$^{th}$ observation and $\hat{y}_i$ is the predicted value of the i$^{th}$ observation.

**Estimation of the panicle grain number**

A different dataset of 198 images of sorghum panicles from those sorghum hybrids was obtained on the same day and using the same equipment, as explained in section Density maps - Data. In this case, two images of the same panicle were taken from opposite sides. The panicles were then harvested and dried at 65$^{o}$C until stable weight. The grains were removed from the panicles and counted using a machine to count grains.

Since a single image cannot represent the total grain number of a panicle, a model was fitted to estimate the total number of grains from the density map estimations. Following the procedures of section 2.2.1, the panicles were manually segmented, the background was removed, and the imagery was patched in three. The output imagery was then input into the best model from section 2.2 (Sorghum-Net), and the number of grains in the image was estimated by the integral of the pixels in the image using the sum function from the numpy library (Harris et al., 2020). The total grains of a single image were set as the sum of the grains of the three patches.

To compare the symmetry of the panicles, the two images of the same panicle on opposite sides were compared using the metrics MAE, MSE, and root mean square error (RMSE) (equation 4), employing the scikit-learn library (Pedregosa et al., 2023). A polynomial linear regression model was fitted using the scikit-learn library (Pedregosa et al., 2023), with the observed counted grains as dependent and the estimation from the density maps as independent, split in a random 70:30% train: test proportion. Finally, to quantify the accuracy of the model, the scikit-learn library (Pedregosa et al., 2023) was employed to calculate the following metrics: coefficient of determination ($R^2$), MSE, RMSE, and MAPE.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (4)$$

# Results

## Detection and Segmentation

The model trained using the Yolov8 framework performed better in all evaluated parameters (Table 1). Yolov8 performed better in correctly classifying a sample as positive (Table 1; AP50, 0.887 and 0.890 vs 0.785 and 0.756), faster (Table 1; inference time, 0.719s vs. 0.925s), and with a smaller model file size (Table 2.1; model weight, 141MB vs. 491MB).

Three examples were chosen to illustrate different scenarios for field conditions (Figure 2.5). Figure 2.5.a shows a partially true positive, with both models (Detectron2 and Yolo8) detecting and segmenting the panicle of interest, but also considering part of the background for the images. Figure 2.5.b shows a partially adequate identification of the panicle for both models; however, the model using Yolov8 presented higher accuracy, considering more parts of the same panicle. Figure 2.5.c shows a situation where was possible to see three panicles in the imagery, and both models identified and segmented them, with the Yolov8 model presenting higher accuracy. Lastly, the model trained using Yolov8 could segment most of the panicles and even better distinguish panicles in the background.

**Density maps**

The model trained using Sorghum-Net architecture presented lower values of MAE, MSE, and MAPE (26, 1271, and 17%, respectively) compared to all the tested models (56, 5350, and 36% for MCNN and 39, 2821, and 28% for TCNN-seed) (Table 2.2). However, MCNN required less computational space because it used a smaller file size of model weights and a smaller number of parameters (1.0MB and 78K, respectively) compared to the other models (1.5MB and 113K for TCNN-Seed and 6.9MB and 568K for Sorghum-Net). Despite these distinctions, no significant differences were found for the inference time among the three models (~16ms), which does not represent a bottleneck for future applications.

All models could successfully estimate the overall density of the image, identifying empty spaces and areas of high grain density (Figure 2.6). The MCNN model overpredicted areas with high grain density, while TCNN-Seed and Sorghum-Net could reasonably estimate grains with high density. Sorghum-Net showed a more diffuse prediction, indicating more uncertainty in predicting the location of the point. This uncertainty was not a negative aspect as it reduced false negatives (over predictions).

### Estimation of the whole panicle

The observed dataset showed a wide range of grain numbers per panicle, with a minimum of 672 to a maximum of 4,156 grains, with a mean of 2,363 grains for the entire dataset.

A polynomial linear model was fitted to predict the total number of grains per panicle using the prediction obtained from the density maps of Sorghum-Net (Figure 2.7.a). This model showed an $R^2$ of 0.48, an MSE of 365122, an RMSE of 604 grains, and a MAPE of 17%. The 1:1 scatter plot showing the observed (ground truth number of grains in the panicles) versus predicted (using the polynomial linear model fitted to the predictions outputted from the density maps) is shown in Figure 2.7.b.

In addition, predictions made using images from opposite sides of the same panicle differed by 103 grains (RMSE). For this task, the MAE was 74 and the MSE was 10524. These results showed that the tested sorghum panicles were rather symmetrical, with minor differences in grain number estimation using one image.

## Discussion

### Framework development

This study developed a framework linking CNN and a linear model to accurately estimate the number of grains in a sorghum panicle using images captured by a smartphone camera under field conditions. Methodologically, our study compares the performance of Yolov8 and Detectron2 for detection and segmentation tasks. Similar comparisons have been conducted in other domains (Ghafari et al., 2022; Prasetyo et al., 2020). Prasetyo et al. (2020)] compared the performance of these models to segment the head and tail of fish with overlapping issues in the images. Similar to our study, Mask-RCNN presented lower precision and recall. The same comparison was made by Ghafari et al. (2022) to detect cells from microfluidic images, where Yolov3 outperformed Mask-RCNN. In these studies, members of the Yolo family outperformed Mask-RCNN, Detectron2's model architecture for segmentation tasks, in terms of precision and recall when segmenting overlapping objects. Likewise other research outcomes (Boominathan et al., 2016; Ghafari et al., 2022; Hu et al., 2022, 2023; Prasetyo et al., 2020), this study also points out that the panicles in the background and foreground panicles share visual similarities, which explains the relatively lower metrics presented in our study, and the overall complexity of unfolding this task under field conditions.

Grain counting, the next step in our framework, presents challenges related to the different scales in the images (Zhang et al., 2016). The panicle images used in this study, similar to the pod images used by Li et al. (Y. Li et al., 2019), presented grains of approximately the same size, which explains the lower metrics achieved by the MCNN model, and the high metrics of TCNN-Seed presented in our study. Sorghum-Net, the model proposed in this work, shows an improved performance, possibly due to its higher complexity, which allows a more effective feature extraction. The metrics and the visual predictions obtained from the density maps showed that the model trained using the Sorghum-Net architecture presented more than half of the error compared to one using the MCNN architecture. Furthermore, the model trained using Sorghum-Net had 1.5 times less error relative to the TCNN-Seed architecture. Therefore, among the tested models, the Sorghum-Net architecture is the best tested option for counting the number of grains within a sorghum panicle using segmented images obtained at the plant level. The size of the weight file and the number of parameters would affect the overall time to complete the predictions depending on the hardware used to perform the computations, but no differences in inference time were observed among the three models (Velasco-Montero et al., 2019; Wang et al., 2018). Lastly, the polynomial linear model predictions based on density models achieved superior accuracy in estimating the number of grains in the entire panicle compared to previous attempts (Ciampitti et al., 2015; Santiago et al., 2023).

**Comparison with existing technologies**

At the plant level, non-invasive attempts to count the number of grains within a sorghum panicle were performed by James et al. (James et al., 2024) using a deep learning method for point cloud feature extraction for grain detection and counting of sorghum panicles. The same authors obtained a mean absolute percentage error of 6.5%; however, the experiment was performed in a laboratory setting using a high quality digital camera and a tent to collect images. Using a smartphone camera at the plant level, Santiago et al. (2023) obtained a root mean square percentage error of 53%, higher than the one obtained in this study. Using a similar approach, but for maize kernels, Khaki et al. (Khaki et al., 2021) developed DeepCorn, a semi-supervised deep learning method, and obtained a mean percent error of 30% and a number of parameters of 26.6 M.

Sorghum high-through phenotyping techniques are also being addressed using UAVs to obtain images. Li et al. (H. Li et al., 2022) compared different deep learning methods to count sorghum panicles; EfficientDet, SSD, and YOLOv4 were used, with the latter showing the best results for the task, with a precision of 98%. Li et al. (Malambo et al., 2019) used deep learning methods to segment sorghum panicles using images obtained from UAVs, with an accuracy of 94%. Despite the aforementioned studies showed excellent results, they did not consider the number of grains within the panicles, which is important information for breeders and yield estimation. UAV and aerial imagery represent the second level of phenotyping, being able to count plants and panicles, not the number of grains.

The use of satellites to predict sorghum yield has also been used. Yang et al. (Yang et al., 2009) employed satellite imagery to estimate crop yield and obtained an $R^2$ of 0.8 in the estimated versus measured yield model. Despite the high $R^2$ obtained, the imagery used had a spatial resolution of 10-m, which did not allow for high-throughput phenotyping at the plant level. Satellite imagery provides a third level of phenotyping, yield estimation at the field or regional level.

Integration of the three levels can increase genetic gain by analyzing different levels of phenotyping and crop aspects (Araus et al., 2018). In the future, satellite or aerial imagery can be used to guide scouting of sections of interest in the field; during scouting, the plant-level handheld or proximal devices will be used at the plant level (Kalischuk et al., 2019). The models presented in this paper provide the cornerstone of these equipment back-ends, counting the number of grains in sorghum panicles, and thus supporting high throughput plant-level phenotyping.

**Limitations & future applications**

The overlap between the background and foreground sorghum panicles presented in the field images led to difficulties in manual labeling and increased false positive values. Nevertheless, it is important to recognize that this limitation is an inherent setting of the proposed framework, as observed in a similar field study (Malambo et al., 2019).

The continuous improvement of these models is essential to increase their precision and reduce the error metrics (Cao et al., 2023). The continuous improvement of these models is essential to increase their precision and reduce the error metrics. Therefore, a future step is to expand the dataset to include images from different lighting conditions, backgrounds, and growth stages. To reduce the effort required to annotate the data, weak supervision training can be used, as proposed by Rühling Cachay et al. and Robinson et al. (Robinson et al., 2020; Rühling Cachay et al., 2021). Performance analysis is fundamental when using CNN models (Véstias, 2019). The size of the weight file and the number of parameters would affect the overall time to complete the predictions depending on the hardware used to perform the computations, but no differences in inference time were observed among the three models (Kristiani et al., 2020; Lane et al., 2016).

The grain number estimation step within a sorghum panicle is a critical component in developing effective yield prediction models (van Oosterom & Hammer, 2008), due to the importance of grain number in yield formation (Ciampitti et al., 2015). The implementation of this approach for smartphone devices will assist farmers and breeders in estimating sorghum yield prior to harvest, a crucial advancement in the development of future digital decision support tools. This application promises to provide timely and reliable sorghum yield information to aid macro-regulation of prices (H. Li et al., 2017; Lipper et al., 2014) and to support with high-throughput phenotyping for breeders, agronomists, and farmers under field conditions (Demarco et al., 2023). Lastly, this framework can be adapted to other crops, such as millet and wheat, to provide a comprehensive approach to estimate yields prior to harvest.

## Conclusion

This study addressed a critical bottleneck in advancing field trait characterization and yield estimation for sorghum crops by integrating computer vision and artificial intelligence with traditional field phenotyping. The proposed framework lays the foundation for the development of a robust plant-level application for yield estimation to benefit farmers and breeders. Further refinement and validation are crucial for broader applicability, and the integration into a plant-level application will be a significant step towards using advanced technologies for efficient sorghum yield estimation.

# Figures

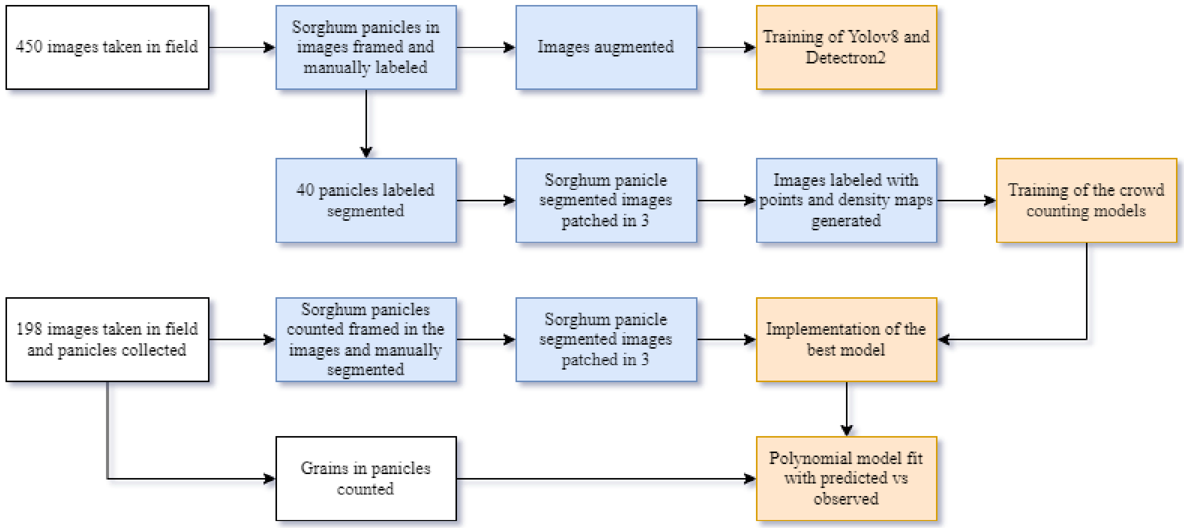


**Figure 2.1. Research diagram of the proposed framework. The white boxes denote the manual component under field conditions, the blue boxes represent the processes executed via the computer, and the yellow boxes refer to the processes linked to machine learning models.**

**Figure 2.2. Photograph showing how the pictures were collected under field conditions.**


**Figure 2.3. Examples of the dataset used in this study: sorghum panicles with different architecture, colors, maturity groups, and panicle density. From top to bottom are images with crescent brightness. In the first column, there are white panicles, in the second one, there are red panicles, in the third there are panicles with different densities. All the images show a different panicle architecture.**

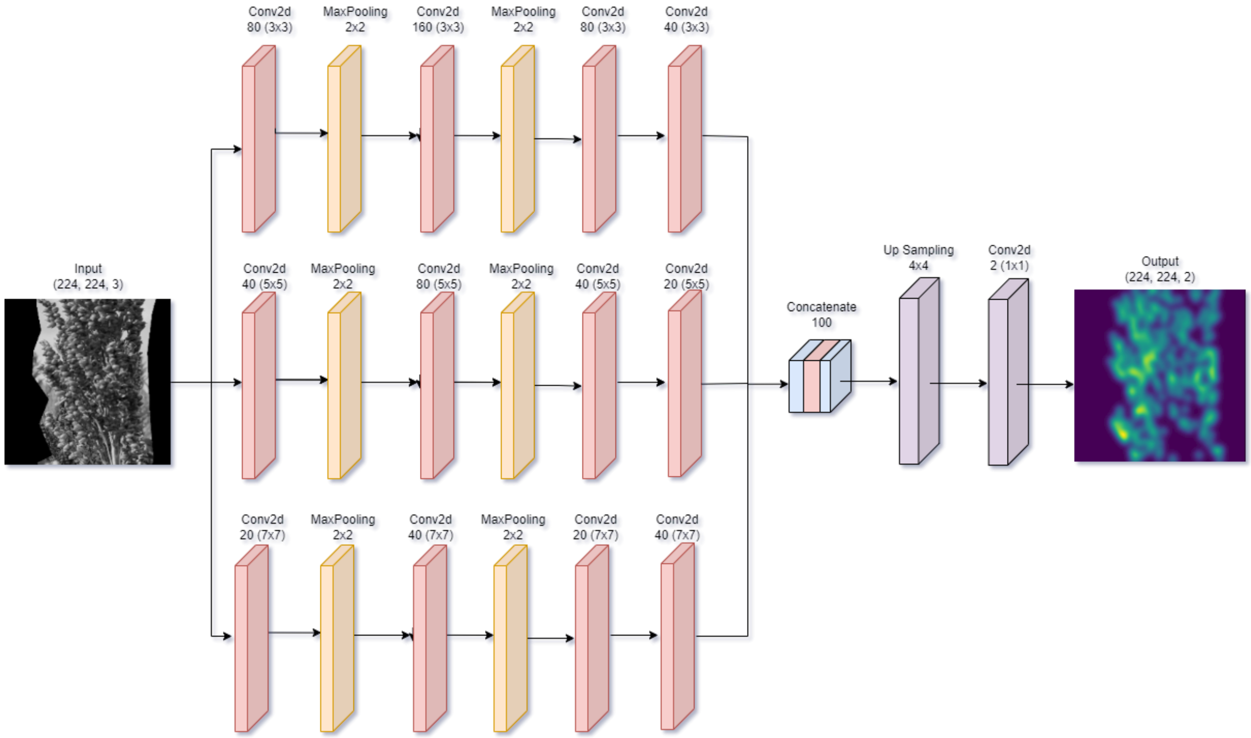


**Figure 2.4. Overview of the Sorghum-Net architecture. The diagram represents the process from the segmented panicle to the density map as the final output, red boxes represent Conv2d, and yellow boxes represent MaxPooling.**

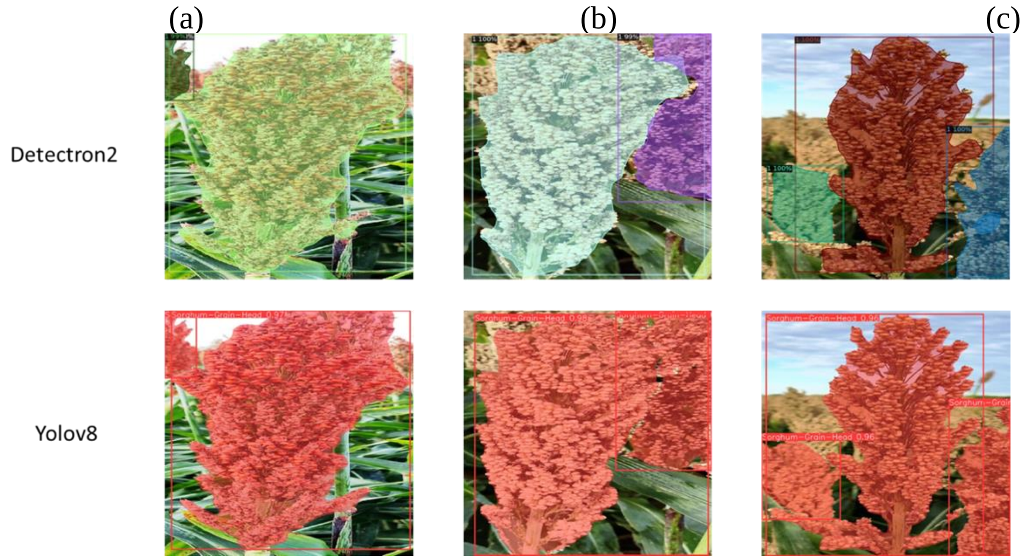


**Figure 2.5. Sorghum panicle segmentation. Predictions were performed by the models trained using the frameworks Detectron2 and Yolov8 over sorghum images**

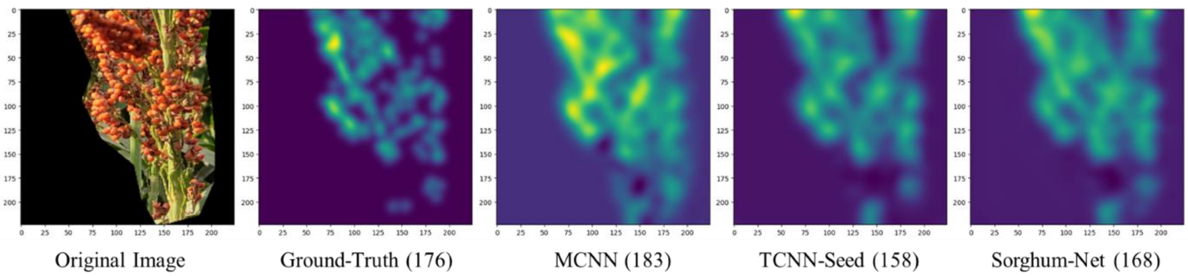


**Figure 2.6. Original image, ground truth, and density map predictions. Original image and density map predictions from the MCNN, TCNN-Seed, and Sorghum-Net models. The number between the parentheses indicates the number of grains present (for the ground truth) and predicted (for the models) in the image.**

# Tables

**Table 2.1. Detection and segmentation models' comparison. Metrics were obtained employing the frameworks Detectron2 and Yolov8 to detect and segment panicles in imagery obtained at field scale using a smartphone camera.**

| Framework | Detection | | Segmentation | | Inference time (s) | Weights file size (MB) |
|---|---|---|---|---|---|---|
| | AP50 | AR | AP50 | AR | | |
| Detectron2 (Meta Research[(*Facebookresearch/Detectron2*, 2019/2023)]) | 0.785 | 0.733 | 0.756 | 0.652 | 0.925 | 491 |
| Yolov8 (Jocher et al. [(Jocher et al., 2022/2023)]) | 0.887 | 0.789 | 0.890 | 0.793 | 0.719 | 141 |

**Table 2.2. Density map models' comparisons. Comparison of the trained models relative to ground truth data to generate density maps from image patches of segmented sorghum panicles taken at field scale using a smartphone camera.**

| Model architecture | MAE | MSE | MAPE | Weights file size (MB) | Number of Parameters |
|---|---|---|---|---|---|
| MCNN (Zhang et al. [(Zhang et al., 2016)]) | 56 | 5350 | 36% | 1.0 | 78K |
| TCNN-Seed (Li et al. [(Y. Li et al., 2019)]) | 39 | 2821 | 28% | 1.5 | 113K |
| Sorghum-Net (This study) | 26 | 1271 | 17% | 6.9 | 568K |

# References


Abstract: Development of a New Tool for Estimating Sorghum Yields at the Farm-Scale. (ASA, CSSA and SSSA International Annual Meetings). (n.d.). Retrieved July 19, 2023, from https://scisoc.confex.com/scisoc/2014am/webprogram/Paper88108.html

Abstract: Sorghumyield® App, a New Tool for Predicting Sorghum Yields: Validation and Calibration. (ASA, CSSA and SSSA International Annual Meetings (2015)). (n.d.). Retrieved March 6, 2024, from https://scisoc.confex.com/crops/2015am/webprogram/Paper93787.html

Albahar, M. (2023). A Survey on Deep Learning and Its Impact on Agriculture: Challenges and Opportunities. Agriculture, 13(3), Article 3. https://doi.org/10.3390/agriculture13030540

Araus, J. L., Kefauver, S. C., Zaman-Allah, M., Olsen, M. S., & Cairns, J. E. (2018). Translating High-Throughput Phenotyping into Genetic Gain. Trends in Plant Science, 23(5), 451–466. https://doi.org/10.1016/j.tplants.2018.02.001

Arya, S., Sandhu, K. S., Singh, J., & kumar, S. (2022). Deep learning: As the new frontier in high-throughput plant phenotyping. Euphytica, 218(4), 47. https://doi.org/10.1007/s10681-022-02992-3

Boominathan, L., Kruthiventi, S. S. S., & Babu, R. V. (2016). CrowdNet: A Deep Convolutional Network for Dense Crowd Counting. Proceedings of the 24th ACM International Conference on Multimedia, 640–644. https://doi.org/10.1145/2964284.2967300

Bradski, G. (2000). The OpenCV Library. Dr. Dobb's Journal: Software Tools for the Professional Programmer, 25(11), 120–123.

C, A., Rao, B. D., & Rao, V. A. T. and T. G. N. (2017). Improving production and utilization of sorghum in Asia Indian Institute of Millets Research, India. In Achieving sustainable cultivation of sorghum Volume 2. Burleigh Dodds Science Publishing.

Cai, Z., & Vasconcelos, N. (2017). Cascade R-CNN: Delving into High Quality Object Detection (arXiv:1712.00726). arXiv. https://doi.org/10.48550/arXiv.1712.00726

Cao, Z., Jiang, K., Zhou, W., Xu, S., Peng, H., & Yang, D. (2023). Continuous improvement of self-driving cars using dynamic confidence-aware reinforcement learning. Nature Machine Intelligence, 5(2), 145–158. https://doi.org/10.1038/s42256-023-00610-y

Cereals & Grains | Inpho | Food and Agriculture Organization of the United Nations. (n.d.). Retrieved July 19, 2023, from https://www.fao.org/in-action/inpho/crop-compendium/cereals-grains/en/

Davies, E. R. (2005). CHAPTER 1—Vision, the Challenge. In E. R. Davies (Ed.), Machine Vision (Third Edition) (pp. 1–15). Morgan Kaufmann. https://doi.org/10.1016/B978-012206093-9/50004-6

Demarco, P. A., Mayor, L., Rotundo, J. L., Prasad, P. V. V., Morris, G. P., Fernandez, J. A., Tamagno, S., Hammer, G., Messina, C. D., & Ciampitti, I. A. (2023). Retrospective study in U.S. commercial sorghum breeding: II. Physiological changes associated to yield gain. Crop Science, 63(2), 867–878. https://doi.org/10.1002/csc2.20845

Deng, J., Dong, W., Socher, R., Li, L.-J., Kai Li, & Li Fei-Fei. (2009). ImageNet: A large-scale hierarchical image database. 2009 IEEE Conference on Computer Vision and Pattern Recognition, 248–255. https://doi.org/10.1109/CVPR.2009.5206848

Facebookresearch/detectron2. (2023). [Python]. Meta Research. https://github.com/facebookresearch/detectron2 (Original work published 2019)

Farjon, G., Huijun, L., & Edan, Y. (2023). Deep-learning-based counting methods, datasets, and applications in agriculture: A review. Precision Agriculture, 24(5), 1683–1711. https://doi.org/10.1007/s11119-023-10034-8

Fountas, S., Carli, G., Sørensen, C. G., Tsiropoulos, Z., Cavalaris, C., Vatsanidou, A., Liakos, B., Canavari, M., Wiebensohn, J., & Tisserye, B. (2015). Farm management information systems: Current situation and future perspectives. Computers and Electronics in Agriculture, 115, 40–50. https://doi.org/10.1016/j.compag.2015.05.011

Ghafari, M., Mailman, D., Hatami, P., Peyton, T., Yang, L., Dang, W., & Qin, H. (2022). A Comparison of YOLO and Mask-RCNN for Detecting Cells from Microfluidic Images. 2022 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), 204–209. https://doi.org/10.1109/ICAIIC54071.2022.9722616

Ghosal, S., Zheng, B., Chapman, S. C., Potgieter, A. B., Jordan, D. R., Wang, X., Singh, A. K., Singh, A., Hirafuji, M., Ninomiya, S., Ganapathysubramanian, B., Sarkar, S., & Guo, W. (2019). A Weakly Supervised Deep Learning Framework for Sorghum Head Detection and Counting. Plant Phenomics, 2019. https://doi.org/10.34133/2019/1525874

Güler, R. A., Neverova, N., & Kokkinos, I. (2018). DensePose: Dense Human Pose Estimation In The Wild (arXiv:1802.00434). arXiv. https://doi.org/10.48550/arXiv.1802.00434

Gupta, P. (2021). APPLICATIONS OF DEEP LEARNING METHODS IN AGRICULTURE: A REVIEW. 8.

Haider, T., Farid, M. S., Mahmood, R., Ilyas, A., Khan, M. H., Haider, S. T.-A., Chaudhry, M. H., & Gul, M. (2021). A Computer-Vision-Based Approach for Nitrogen Content Estimation in Plant Leaves. Agriculture, 11(8), Article 8. https://doi.org/10.3390/agriculture11080766

Hariprasanna, K., & Rakshit, S. (2016). Economic Importance of Sorghum. In S. Rakshit & Y.-H. Wang (Eds.), The Sorghum Genome (pp. 1–25). Springer International Publishing. https://doi.org/10.1007/978-3-319-47789-3_1

Harris, C. R., Millman, K. J., Van Der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., Van Kerkwijk, M. H., Brett, M., Haldane, A., Del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. Nature, 585(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2018). Mask R-CNN (arXiv:1703.06870). arXiv. https://doi.org/10.48550/arXiv.1703.06870

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition (arXiv:1512.03385). arXiv. https://doi.org/10.48550/arXiv.1512.03385

Hu, H., Kaizu, Y., Zhang, H., Xu, Y., Imou, K., Li, M., Huang, J., & Dai, S. (2022). Recognition and localization of strawberries from 3D binocular cameras for a strawberry picking robot using coupled YOLO/Mask R-CNN. International Journal of Agricultural and Biological Engineering, 15(6), Article 6. https://doi.org/10.25165/ijabe.v15i6.7306

Hu, H., Tang, C., Shi, C., & Qian, Y. (2023). Detection of residual feed in aquaculture using YOLO and Mask RCNN. Aquacultural Engineering, 100, 102304. https://doi.org/10.1016/j.aquaeng.2022.102304

James, C., Smith, D., He, W., Chandra, S. S., & Chapman, S. C. (2024). GrainPointNet: A deep-learning framework for non-invasive sorghum panicle grain count phenotyping. Computers and Electronics in Agriculture, 217, 108485. https://doi.org/10.1016/j.compag.2023.108485

Jocher, G., Chaurasia, A., & Qiu, J. (2023). YOLO by Ultralytics (Version 8.0.0) [Python]. https://github.com/ultralytics/ultralytics (Original work published 2022)

Jung, A. (2023). Imgaug [Python]. https://github.com/aleju/imgaug (Original work published 2015)

Kalischuk, M., Paret, M. L., Freeman, J. H., Raj, D., Da Silva, S., Eubanks, S., Wiggins, D. J., Lollar, M., Marois, J. J., Mellinger, H. C., & Das, J. (2019). An Improved Crop Scouting Technique Incorporating Unmanned Aerial Vehicle–Assisted Multispectral Crop Imaging into Conventional Scouting Practice for Gummy Stem Blight in Watermelon. Plant Disease, 103(7), 1642–1650. https://doi.org/10.1094/PDIS-08-18-1373-RE

Khaki, S., Pham, H., Han, Y., Kuhl, A., Kent, W., & Wang, L. (2021). DeepCorn: A semi-supervised deep learning method for high-throughput image-based corn kernel counting and yield estimation. Knowledge-Based Systems, 218, 106874. https://doi.org/10.1016/j.knosys.2021.106874

Kristiani, E., Yang, C.-T., & Phuong Nguyen, K. L. (2020). Optimization of Deep Learning Inference on Edge Devices. 2020 International Conference on Pervasive Artificial Intelligence (ICPAI), 264–267. https://doi.org/10.1109/ICPAI51961.2020.00056

Lane, N. D., Bhattacharya, S., Georgiev, P., Forlivesi, C., Jiao, L., Qendro, L., & Kawsar, F. (2016). DeepX: A Software Accelerator for Low-Power Deep Learning Inference on Mobile Devices. 2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN), 1–12. https://doi.org/10.1109/IPSN.2016.7460664

Li, H., Chen, Z., Liu, G., Jiang, Z., & Huang, C. (2017). Improving Winter Wheat Yield Estimation from the CERES-Wheat Model to Assimilate Leaf Area Index with Different Assimilation Methods and Spatio-Temporal Scales. Remote Sensing, 9(3), Article 3. https://doi.org/10.3390/rs9030190

Li, H., Wang, P., & Huang, C. (2022). Comparison of Deep Learning Methods for Detecting and Counting Sorghum Heads in UAV Imagery. Remote Sensing, 14(13), Article 13. https://doi.org/10.3390/rs14133143

Li, M., Shao, M.-R., Zeng, D., Ju, T., Kellogg, E. A., & Topp, C. N. (2020). Comprehensive 3D phenotyping reveals continuous morphological variation across genetically diverse sorghum inflorescences. New Phytologist, 226(6), 1873–1885. https://doi.org/10.1111/nph.16533

Li, Y., Jia, J., Zhang, L., Khattak, A. M., Sun, S., Gao, W., & Wang, M. (2019). Soybean Seed Counting Based on Pod Image Using Two-Column Convolution Neural Network. IEEE Access, 7, 64177–64185. https://doi.org/10.1109/ACCESS.2019.2916931

Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2018). Focal Loss for Dense Object Detection (arXiv:1708.02002). arXiv. https://doi.org/10.48550/arXiv.1708.02002

Lin, T.-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., & Dollár, P. (2015). Microsoft COCO: Common Objects in Context (arXiv:1405.0312). arXiv. https://doi.org/10.48550/arXiv.1405.0312

Lin, Z., & Guo, W. (2020). Sorghum Panicle Detection and Counting Using Unmanned Aerial System Images and Deep Learning. Frontiers in Plant Science, 11. https://www.frontiersin.org/articles/10.3389/fpls.2020.534853

Lipper, L., Thornton, P., Campbell, B. M., Baedeker, T., Braimoh, A., Bwalya, M., Caron, P., Cattaneo, A., Garrity, D., Henry, K., Hottle, R., Jackson, L., Jarvis, A., Kossam, F., Mann, W., McCarthy, N., Meybeck, A., Neufeldt, H., Remington, T., … Torquebiau, E. F. (2014). Climate-smart agriculture for food security. Nature Climate Change, 4(12), Article 12. https://doi.org/10.1038/nclimate2437

Maheswari, P., Raja, P., Apolo-Apolo, O. E., & Pérez-Ruiz, M. (2021). Intelligent Fruit Yield Estimation for Orchards Using Deep Learning Based Semantic Segmentation Techniques—A Review. Frontiers in Plant Science, 12. https://doi.org/10.3389/fpls.2021.684328

Malambo, L., Popescu, S., Ku, N.-W., Rooney, W., Zhou, T., & Moore, S. (2019). A Deep Learning Semantic Segmentation-Based Approach for Field-Level Sorghum Panicle Counting. Remote Sensing, 11(24), Article 24. https://doi.org/10.3390/rs11242939

Mendes, J., Pinho, T. M., Neves dos Santos, F., Sousa, J. J., Peres, E., Boaventura-Cunha, J., Cunha, M., & Morais, R. (2020). Smartphone Applications Targeting Precision Agriculture Practices—A Systematic Review. Agronomy, 10(6), Article 6. https://doi.org/10.3390/agronomy10060855

Monk, R. L., Miller, F. R., & McBee, G. G. (1984). Sorghum improvement for energy production. Biomass, 6(1), 145–153. https://doi.org/10.1016/0144-4565(84)90017-9

Oh, M., Olsen, P., & Ramamurthy, K. N. (2019). Counting and Segmenting Sorghum Heads (arXiv:1905.13291). arXiv. https://doi.org/10.48550/arXiv.1905.13291

Patel, A. K., Park, E.-S., Lee, H., Priya, G. G. L., Kim, H., Joshi, R., Arief, M. A. A., Kim, M. S., Baek, I., & Cho, B.-K. (2023). Deep Learning-Based Plant Organ Segmentation and Phenotyping of Sorghum Plants Using LiDAR Point Cloud. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 16, 8492–8507. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. https://doi.org/10.1109/JSTARS.2023.3312815

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., & Cournapeau, D. (n.d.). Scikit-learn: Machine Learning in Python. MACHINE LEARNING IN PYTHON.

Petrellis, N. (2017). A smart phone image processing application for plant disease diagnosis. 2017 6th International Conference on Modern Circuits and Systems Technologies (MOCAST), 1–4. https://doi.org/10.1109/MOCAST.2017.7937683

Prasetyo, E., Suciati, N., & Fatichah, C. (2020). A Comparison of YOLO and Mask R-CNN for Segmenting Head and Tail of Fish. 2020 4th International Conference on Informatics and Computational Sciences (ICICoS), 1–6. https://doi.org/10.1109/ICICoS51170.2020.9299024

Raj, H. C., Boomiraj, K., Pazhanivelan, S., Moorthy, N. K. S., & Prasanthrajan, M. (2023). Spatial Yield Assessment of Sorghum (Sorghum Bicolor) Using DSSAT – CERES – Sorghum Model in Solapur, Maharashtra, India. International Journal of Environment and Climate Change, 13(10), 546–554. https://doi.org/10.9734/ijecc/2023/v13i102684

Robinson, J., Jegelka, S., & Sra, S. (2020). Strength from Weakness: Fast Learning Using Weak Supervision. Proceedings of the 37th International Conference on Machine Learning, 8127–8136. https://proceedings.mlr.press/v119/robinson20a.html

Roboflow: Give your software the power to see objects in images and video. (n.d.). Retrieved July 25, 2023, from https://roboflow.com/

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation (arXiv:1505.04597). arXiv. https://doi.org/10.48550/arXiv.1505.04597

Rühling Cachay, S., Boecking, B., & Dubrawski, A. (2021). End-to-End Weak Supervision. Advances in Neural Information Processing Systems, 34, 1845–1857. https://proceedings.neurips.cc/paper_files/paper/2021/hash/0e674a918ebca3f78bfe02e2f387689d-Abstract.html

Santiago, G., Carcedo, A., Marziotte, L., & Ciampitti, I. (2023). Algorithm to Estimate Sorghum Grain Number from Panicles Using Images Collected with a Smartphone at Field-Scale. Kansas Agricultural Experiment Station Research Reports, 9(4). https://doi.org/10.4148/2378-5977.8465

Santos, L., Santos, F. N., Oliveira, P. M., & Shinde, P. (2020). Deep Learning Applications in Agriculture: A Short Review. In M. F. Silva, J. Luís Lima, L. P. Reis, A. Sanfeliu, & D. Tardioli (Eds.), Robot 2019: Fourth Iberian Robotics Conference (pp. 139–151). Springer International Publishing. https://doi.org/10.1007/978-3-030-35990-4_12

Srinivasa Rao, P., Reddy, B. V. S., Nagaraj, N., & Upadhyaya, H. D. (2014). Sorghum Production for Diversified Uses (Y. H. Wang, H. D. Upadhyaya, & C. Kole, Eds.; pp. 01–27). CRC Press (Taylor & Francis). http://www.crcnetbase.com/doi/abs/10.1201/b17153-2

van Oosterom, E. J., & Hammer, G. L. (2008). Determination of grain number in sorghum. Field Crops Research, 108(3), 259–268. https://doi.org/10.1016/j.fcr.2008.06.001

Velasco-Montero, D., Femández-Bemi, J., Carmona-Gálán, R., & Rodríguez-Vázquez, A. (2019). On the Correlation of CNN Performance and Hardware Metrics for Visual Inference on a Low-Cost CPU-based Platform. 2019 International Conference on Systems, Signals and Image Processing (IWSSIP), 249–254. https://doi.org/10.1109/IWSSIP.2019.8787329

Véstias, M. P. (2019). A Survey of Convolutional Neural Networks on Edge with Reconfigurable Computing. Algorithms, 12(8), Article 8. https://doi.org/10.3390/a12080154

Vieyra, R. E., Megowan-Romanowicz, C., O'Brien, D. J., Vieyra, C., & Johnson-Glenberg, M. C. (2023). Harnessing the Digital Science Education Revolution: Smartphone Sensors as

Teaching Tools. In Theoretical and Practical Teaching Strategies for K-12 Science Education in the Digital Age (pp. 131–152). IGI Global. https://doi.org/10.4018/978-1-6684-5585-2.ch008

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., Van Der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … Vázquez-Baeza, Y. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. Nature Methods, 17(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2

Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep Learning for Computer Vision: A Brief Review. Computational Intelligence and Neuroscience, 2018, e7068349. https://doi.org/10.1155/2018/7068349

Wan, S., & Goudos, S. (2020). Faster R-CNN for multi-class fruit detection using a robotic vision system. Computer Networks, 168, 107036. https://doi.org/10.1016/j.comnet.2019.107036

Wang, J., Lin, J., & Wang, Z. (2018). Efficient Hardware Architectures for Deep Convolutional Neural Network. IEEE Transactions on Circuits and Systems I: Regular Papers, 65(6), 1941–1953. IEEE Transactions on Circuits and Systems I: Regular Papers. https://doi.org/10.1109/TCSI.2017.2767204

Wiley, V., & Lucas, T. (2018). Computer Vision and Image Processing: A Paper Review. International Journal of Artificial Intelligence Research, 2(1), Article 1. https://doi.org/10.29099/ijair.v2i1.42

Yang, C., Everitt, J. H., & Bradford, J. M. (2009). Evaluating high resolution SPOT 5 satellite imagery to estimate crop yield. Precision Agriculture, 10(4), 292–303. https://doi.org/10.1007/s11119-009-9120-6

Ye, X., Abe, S., Zhang, S., & Yoshimura, H. (2020). Rapid and non-destructive assessment of nutritional status in apple trees using a new smartphone-based wireless crop scanner system. Computers and Electronics in Agriculture, 173, 105417. https://doi.org/10.1016/j.compag.2020.105417

Young, S. N., Kayacan, E., & Peschel, J. M. (2019). Design and field evaluation of a ground robot for high-throughput phenotyping of energy sorghum. Precision Agriculture, 20(4), 697–722. https://doi.org/10.1007/s11119-018-9601-6

Zannou, J. G. N., & Houndji, V. R. (2019). Sorghum Yield Prediction using Machine Learning. 2019 3rd International Conference on Bio-Engineering for Smart Technologies (BioSMART), 1–4. https://doi.org/10.1109/BIOSMART.2019.8734219

Zhang, Y., Zhou, D., Chen, S., Gao, S., & Ma, Y. (2016). Single-Image Crowd Counting via Multi-Column Convolutional Neural Network. 589–597. https://openaccess.thecvf.com/content_cvpr_2016/html/Zhang_Single-Image_Crowd_Counting_CVPR_2016_paper.html

# Chapter 3 – A mobile application for sorghum panicle photometry and yield estimation


## Abstract

Estimating sorghum yield before harvest is laborious and time-consuming, requiring panicle (head) removal and manual or machine counting of grains. Employing smartphones and computer vision algorithms represents a potential solution to this bottleneck. Following this rationale, this application note aims to develop a mobile application that counts the number of grains within a sorghum panicle using imagery tested at both lab and field settings. The application can provide an estimate of yields based on inputs from users and generate a final report. Two deep learning and one linear model from previous work were employed in a Flutter framework-based mobile application to estimate the number of grains from imagery obtained from sorghum panicles at the field scale. Two branches were developed for this application: one local employing the models in the smartphone hardware and one remote employing the models in a web server. Validation for final number of grains was performed in the laboratory and in the field, resulting in a mean average precision error of 22.4% and 21.4%, respectively. The mobile application presented in this study can be used by farmers, crop insurance companies, agronomists, and breeders to rapidly estimate grain numbers in sorghum panicle and forecast yields before harvest.

## Introduction

Estimating yield before harvest is a challenging but crucial undertaking that allows farmers to organize their post-harvest and marketing operations. However, a more precise yield estimation for sorghum is laborious and destructive, relying on the number of panicles per unit of area, the number of grains within a panicle, and the grain mass (Ciampitti & Carcedo, 2023). From the three steps mentioned before, the key for estimating sorghum yield is the last: counting the number of grains within panicles, as the main yield component associated to yield (Ciampitti et al., 2019). However, there is a bottleneck in this step: panicle extraction and human or machine counting are laborious, cost-intensive, and time-consuming. This process requires of a destructive approach to extract panicles, adding more complexity to this challenging process for yield estimation.

An affordable and non-destructive way to address the above bottleneck is to use mobile applications. Examples of mobile applications developed to aid in counting agricultural products and estimating yield can be found in the current scientific literature. For example, Hernández-Hernández et al. (2017) developed a mobile application to count the number of plants in horticultural crops using image color analysis. Aquino et al. (2018) developed vitisBerry, which employs a neural network to estimate the number of berries within clusters in vineyards. Ciampitti et al. (2015) developed a first attempt to determine the number of grains in a sorghum crop by collecting images of a panicle but using a standard square of known size in the background to estimate the volume of the harvestable organ. However, to the best of our knowledge, a mobile app has not yet been developed that can count the grains of a sorghum panicle using image data to predict on-farm crop yield prior to harvest.

Following this rationale, this application note aims to develop a mobile app that counts the number of grains within a sorghum panicle using field scale imagery, herein termed sorghum panicle photometry. The application can assist in forecasting yields using user inputs and the output obtained from sorghum images (grain numbers) to generate a report for the user with sorghum yield estimation.

## Materials and Methods (App development)

### Front-End

The Flutter framework (Flutter, 2024) was used to develop the Front-End application using Dart computing language (Bracha, 2015). The Flutter framework was chosen due to its applications on multi-operational systems. In addition, the Flutter framework is a well-documented and widely used among all the mobile applications, reducing the difficulty of developing and maintaining the application.

The application presents a straightforward interface and user experience, clarifying the steps to use the app. Firstly, the user must agree with the terms of service, this is a relevant step since our mobile application will also store data for scientific purposes. On the main screen, the user selects whether to the “Panicle Photometry” option to estimate grain number or the option to “Estimate Yield”. Both options follow the same steps, but the “Panicle Photometry” option does not include user field input or generate a report. After clicking the “Estimate Yield” button, the user must insert field crop management information. In the next step, the user selects up to 5 images from the photo gallery of the smartphone device or the user can obtain those images with the camera. The images are processed, panicles are segmented, and the segmented panicles are returned to the application screen. The user selects the desired panicles, and the total number of grains is estimated. On the next page, the user obtains the estimated number of grains for each panicle. In the penultimate step, the user enters the seed weight (estimation before harvest), and the forecasted yield is calculated. Finally, a portable document format (in PDF format) report is generated, and the data is saved. Figure 3.1 provides a graphical representation of the application workflow, summarizing the above steps.

The application design, user interface, warning texts, and other functionalities were enhanced by feedback provided by scholars, faculty, and farmers. A video exemplifying the usage, and the application interface is provided in Supplementary Material. Figure 3.2 also provides some screenshots of the application developed in this study, simulating a yield estimation.

**Back End**

The cornerstone of this application is a combination of three different models: two deep neural networks and one linear model, as described by Santiago et al. (2024). The same authors documented that the detection and segmentation of the panicles performed by a Yolov8 with a dataset of 4500 images, presented an average precision of 89% for segmentation. To count the number of grains within a segmented sorghum panicle image, a multicolumn neural network, Sorghum-Net (Santiago et al., 2024). Sorghum-Net was trained with a dataset of 4800 images, presenting a mean average precision error (MAPE) of 17%. Lastly, a polynomial linear model was fit to a dataset of 198 images from 99 manually grain-counted panicles, presenting a MAPE of 17%. For more information regarding the datasets, models, training, and results obtained, please visit (Santiago et al., 2024).

Two branches of the same application were developed. One, using a local computer (smartphone hardware) and aiming to be independent of the Internet to perform the inference.

The second one, using a remote server with a graphics processing unit (GPU) acceleration to reduce the inference time performance while relying on an Internet connection.

The back-end functions of the local branch were developed using Dart computer language (Bracha, 2015). The segmentation and density map models were converted to the tflite file format (Abadi et al., 2016) for use in this branch. The “flutter_vision” package (vladimir, 2022/2024) was used to segment the images using the Yolov8 model, and the “tflite_flutter” package (Tensorflow/Flutter-Tflite, 2023/2024) to derive the density maps inferences. The back-end functions in the remote branch were developed using the Flask framework (Dwyer et al., 2017) and the Python computing language (Van Rossum & Drake, 2009) and hosted on a web server using GPU acceleration. The “ultralytics” (Ultralytics, 2024) package was used to segment the images using the Yolov8 model, and the “TensorFlow” (Abadi et al., 2016) package was employed to obtain the density map inferences. In this branch, when the user clicks on Generate a Report, in addition to the PDF report generation, data is stored in a remote non-relational (NoSQL) database for scientific purposes. The stored data is the same as that shown to the user in the report: field name, hybrid analyzed, seeding rate, field size, number of heads in a unit area, average number of grains counted by the application, row size, estimated yield, seed weight entered by the user, metric system (metric or imperial), and location.

## Validation

Three tests were performed to validate the application: i) laboratory test to verify the overall performance of the mobile app in a controlled environment, ii) a field test to verify the performance of the app using images and data collected at a field scale, iii) a yield estimation using the same panicles from the field test. The dataset used in the laboratory test consisted of 114 images of sorghum panicles obtained using an ordinary smartphone (Xiaomi Poco X3) with a white sheet behind the panicles and controlled illumination. The number of grains within each panicle was manually counted using an optical machine (801 Count-a-Pak from Seedburo). To obtain the predicted number from the application usage, the option “Panicle Photometry” was selected, and the images were imported from the gallery in batches of five images at a time. The predicted numbers of grains were manually annotated. For the field setting we used a field trial with 4 different genotypes. We selected five representative panicles from a row and counted the number of panicles in five-meter length. Then, we framed the interest panicle, parallel to the camera and captured the image. In the end, we captured 80 pictures from the same number of panicles. The panicles were then harvested, threshed and the grains counted using an optical machine (801 Count-a-Pak from Seedburo). To obtain the observed yield values, in addition to the number of grains, the grains from each individual panicle were also weighted to obtain the number of grains in a unit of mass. Using the option “Estimate Yield”, we input the information recorded, among the pictures, generating and estimation of yield per row. The number of panicle population and number of grains per unit of mass values were the same as those obtained from the observed data. Lastly, the data was stored in the NoSQL cloud database and retrieved to evaluate the metrics.

The root mean square error (RMSE), the relative root mean square error (RRMSE), and the Kling-Gupta efficiency (KGE) metrics were used to compare the observed and predicted results obtained in this validation step. The R computing language (R Core Team, 2024) and the library metrica (Correndo et al., 2023) were used to obtain the metrics, and ggplot2 (Wickham et al., 2024) was used to obtain the observed versus predicted graphs. The RMSE for the laboratory validation was 350 grains, representing 22.4% of RRMSE and a KGE value of 0.8. For the field validation, the RMSE was equal to 423 grains, representing 24.4% of RRMSE and presenting a

KGE of 0.45. Lastly, the yield validation presented a RRMSE of 1291 kg/ha, a RRMSE of 15.1%, and a KGE of 0.76 (Table 3.1). Despite the RRMSE of the laboratory validation are higher than the field, the observed versus predicted plot for in-laboratory validation (Figure 3.3a) is more aggrouped, in contrast with the field validation plot, which is more dispersed (Figure 3.3b).

While testing the app, we observed that images with dimensions smaller than the ones obtained from a smartphone camera present significant difference compared to the observed values.

## Discussion and Outlook

This paper presents a mobile application developed to estimate the number of grains in sorghum panicles using images obtained by the mobile device. In addition to the estimation, the application estimates grain yield using user input. The storage of the data input by the users in a remote database provides the cornerstone for future studies regarding grain yield potential, hybrid usage, plant density, and other geolocated information related to sorghum crops. These data alone, and/or combined with other data such as weather, social, and economic data, can fill a recurrent data gap encountered in many research studies (Kebede et al., 2024).

Both branches developed present positive and negative aspects. The local branch allows its usage regarding the internet connection, a common issue in many agricultural fields worldwide (Mack et al., 2024; Villapol et al., 2018). However, it is highly dependent on the smartphone hardware quality to perform inferences with high performance (Li et al., 2024). The inference performance issue in the local branch can be overcome using the remote branch due to the GPU employment (Hanindhito & John, 2024). In addition to improving the inference time, the remote branch also allows SQL data storage in the cloud. The limitation of the remote branch is related to the dependency on internet connection and speed for data processing since images must be uploaded to the web server and downloaded back to the smartphone. A limitation for both applications is related to the image size.

The mobile application can be used by farmers, crop insurance companies, agronomists, and breeders to rapidly assess grain numbers and obtain a grain yield estimation before harvest. These assessments aid in high-throughput phenotyping under field conditions (Araus et al., 2018) and in the determination of on-farm yield forecasts (Chitsiko et al., 2022). Future steps can include deploying the remote branch since a web server must be maintained, developing accounts for the historic archive, and adding grain estimations from other crops within the same application.

# Figures

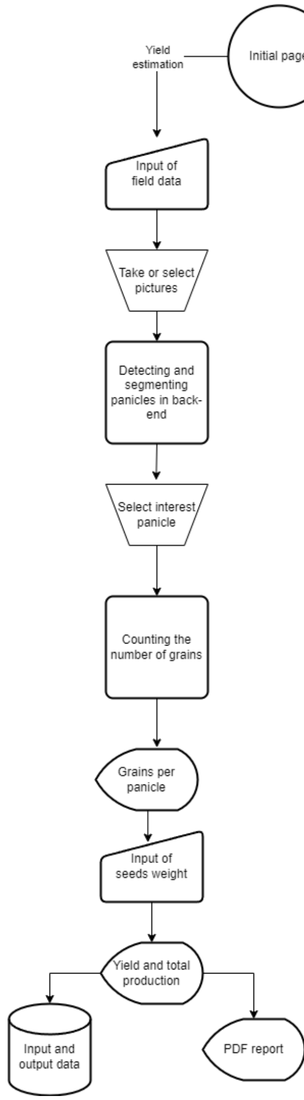


**Figure 3.1. Diagram of the application workflow, where the circle represents the starting point, the rounded corner trapeze represents the user inputs, the sharp corner represents a user action, the rectangles represent the processes executed by the application, the cylinder represents the database, and the ellipsoid represents an output from the application**

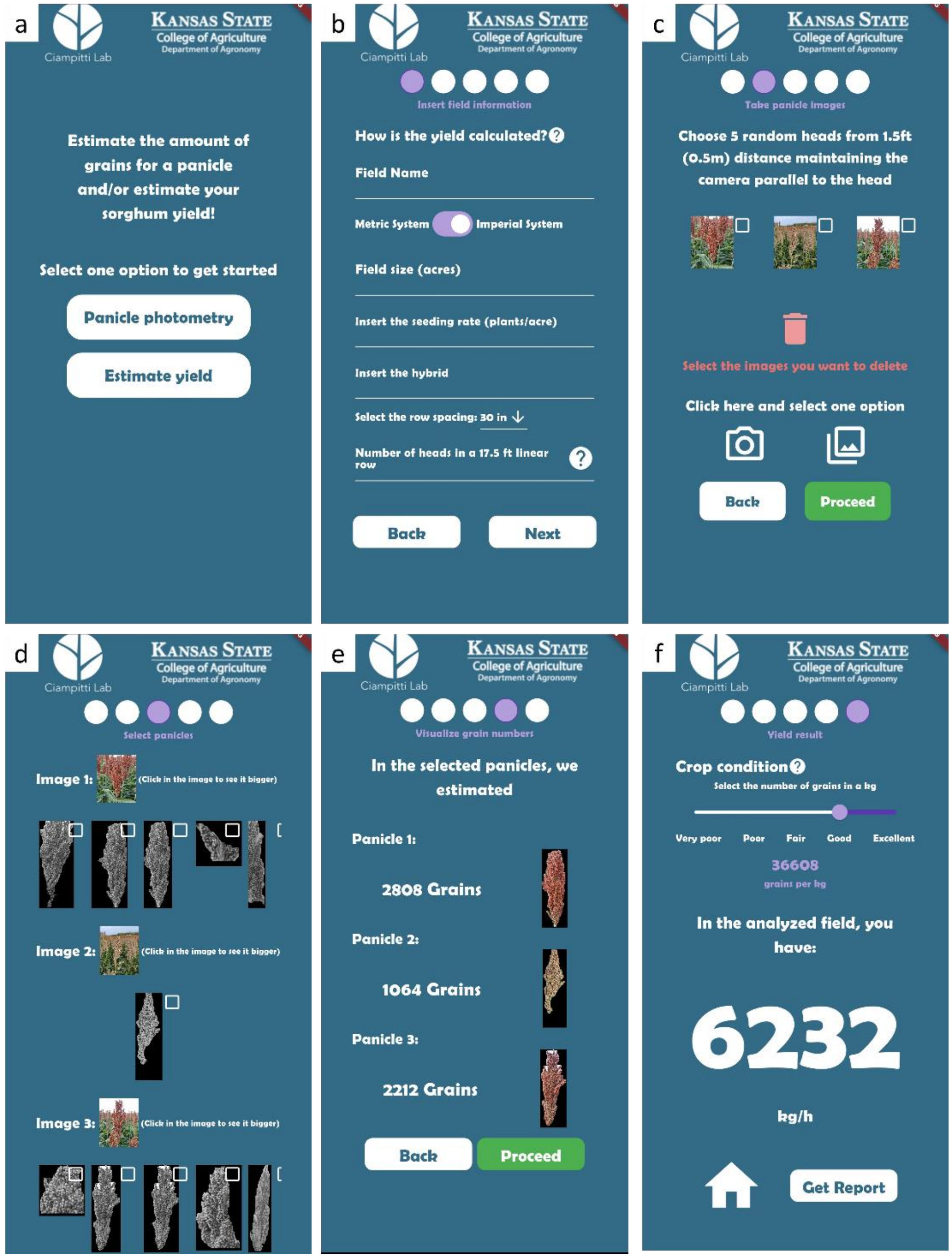


**Figure 3.2. Screenshots of the mobile application developed in this study. It includes the initial page (a), input of field data (b), selection of capture of image (c), selection of segmented panicles (d), exhibition of counted number of grains (e), and crop condition selection and predicted yield output (f)**

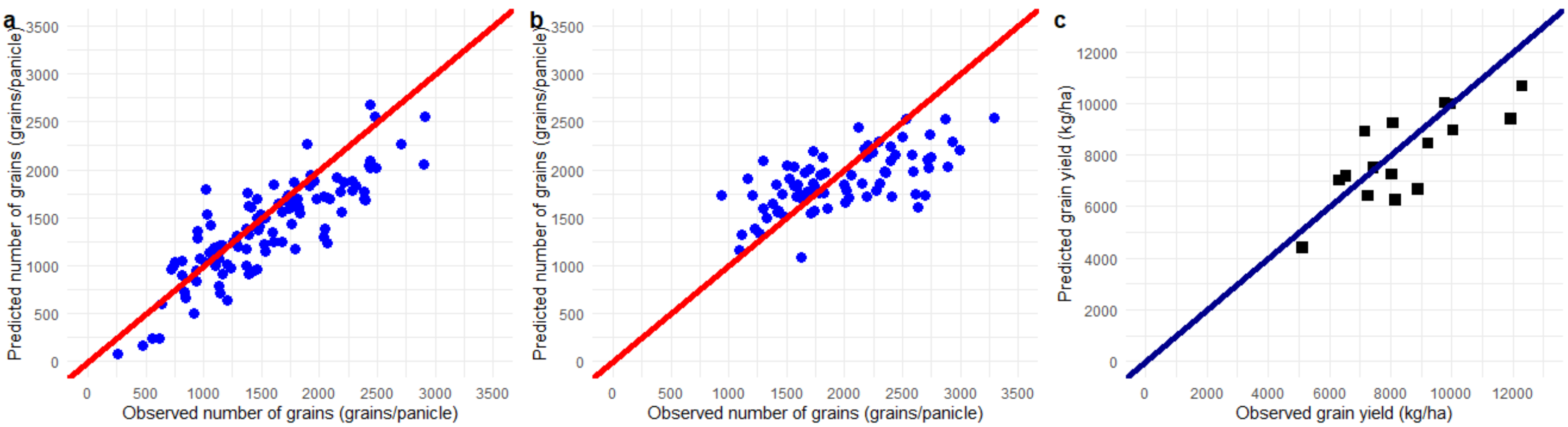


**Figure 3.3. Observed versus Predicted scatter plot obtained from laboratory (a), field (b) both for grain numbers in sorghum panicle, and yield (c) at field scale (grain numbers derived from panel b, using a constant for grain weight and plant density to convert to area, hectare) during the validation of the application.**

# Tables

**Table 3.1. Metrics obtained while evaluating the application presented in this study targeting laboratory for both grain numbers in sorghum panicle and grain yield estimations at field-scale.**

| Validation target | Metrics | | |
|---|---|---|---|
| | RMSE | RRMSE (%) | KGE |
| Laboratory | 350 grains | 22.4 | 0.8 |
| Field | 423 grains | 21.4 | 0.45 |
| Grain Yield | 1291 kg/ha | 15.1 | 0.76 |

# References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., … Zheng, X. (n.d.). *TensorFlow: A system for large-scale machine learning*.

Aquino, A., Barrio, I., Diago, M.-P., Millan, B., & Tardaguila, J. (2018). vitisBerry: An Android-smartphone application to early evaluate the number of grapevine berries by means of image analysis. *Computers and Electronics in Agriculture*, *148*, 19–28. https://doi.org/10.1016/j.compag.2018.02.021

Araus, J. L., Kefauver, S. C., Zaman-Allah, M., Olsen, M. S., & Cairns, J. E. (2018). Translating High-Throughput Phenotyping into Genetic Gain. *Trends in Plant Science*, *23*(5), 451–466. https://doi.org/10.1016/j.tplants.2018.02.001

Bracha, G. (2015). *The Dart Programming Language* (1st ed.). Addison-Wesley Professional.

Chitsiko, R. J., Mutanga, O., Dube, T., & Kutywayo, D. (2022). Review of current models and approaches used for maize crop yield forecasting in sub-Saharan Africa and their potential use in early warning systems. *Physics and Chemistry of the Earth, Parts A/B/C*, *127*, 103199. https://doi.org/10.1016/j.pce.2022.103199

Ciampitti, I. A., Prasad, P. v. v., Schlegel, A. J., Haag, L., Schnell, R. W., Arnall, B., & Lofton, J. (2019). Genotype × Environment × Management Interactions: US Sorghum Cropping Systems. In *Sorghum* (pp. 277–296). John Wiley & Sons, Ltd. https://doi.org/10.2134/agronmonogr58.c13

Ciampitti IA, Azevedo AJ, Balboa GR, Bossio N. 2015. Sorghum yield app, a new tool for predicting sorghum yields: validation and calibration. ASA, CSSA, and SSSA International Annual Meeting (2015). Poster Number: 501. https://scisoc.confex.com/crops/2015am/webprogram/ Paper93787.html

Ciampitti I.A., Carcedo, A. J. P. *Agronomy eUpdate September 14th, 2023: Issue 974*. (n.d.). Retrieved September 4, 2024, from https://eupdate.agronomy.ksu.edu/eu_article_prep.php?article_id=3560

Correndo, A. A., Correndo, A. A., Rosso, L. H. M., Schwalbert, R., Hernandez, C., Bastos, L. M., Nieto, L., Holzworth, D., & Ciampitti, I. A. (2023). *metrica: Prediction Performance Metrics* (Version 2.0.3) [Computer software]. https://cran.r-project.org/web/packages/metrica/index.html

Dwyer, G., Aggarwal, S., & Stouffer, J. (2017). *Flask: Building Python Web Services*. Packt Publishing.

*Flutter documentation*. (n.d.). Retrieved September 4, 2024, from https://docs.flutter.dev

Hanindhito, B., & John, L. K. (2024). Accelerating ML Workloads using GPU Tensor Cores: The Good, the Bad, and the Ugly. *Proceedings of the 15th ACM/SPEC International Conference on Performance Engineering*, 178–189. https://doi.org/10.1145/3629526.3653835

Hernández-Hernández, J. L., Ruiz-Hernández, J., García-Mateos, G., González-Esquiva, J. M., Ruiz-Canales, A., & Molina-Martínez, J. M. (2017). A new portable application for automatic segmentation of plants in agriculture. *Agricultural Water Management*, *183*, 146–157. https://doi.org/10.1016/j.agwat.2016.08.013

Kebede, E. A., Abou Ali, H., Clavelle, T., Froehlich, H. E., Gephart, J. A., Hartman, S., Herrero, M., Kerner, H., Mehta, P., Nakalembe, C., Ray, D. K., Siebert, S., Thornton, P., & Davis, K. F. (2024). Assessing and addressing the global state of food production data scarcity. *Nature Reviews Earth & Environment*, *5*(4), 295–311. https://doi.org/10.1038/s43017-024-00516-2

Li, Z., Paolieri, M., & Golubchik, L. (2024). Inference latency prediction for CNNs on heterogeneous mobile devices and ML frameworks. *Performance Evaluation*, *165*, 102429. https://doi.org/10.1016/j.peva.2024.102429

Mack, E. A., Loveridge, S., Keene, T., & Mann, J. (2024). A Review of the Literature About Broadband Internet Connections and Rural Development (1995-2022). *International Regional Science Review*, *47*(3), 231–292. https://doi.org/10.1177/01600176231202457

*R: The R Project for Statistical Computing*. (n.d.). Retrieved August 1, 2024, from https://www.r-project.org/

Santiago, G. N., Cisdeli Magalhaes, P. H., Carcedo, A. J. P., Marziotte, L., Mayor, L., & Ciampitti, I. A. (2024). Deep Learning Methods Using Imagery from a Smartphone for Recognizing Sorghum Panicles and Counting Grains at a Plant Level. *Plant Phenomics*, *6*, 0234. https://doi.org/10.34133/plantphenomics.0234

*Tensorflow/flutter-tflite*. (2024). [Dart]. tensorflow. https://github.com/tensorflow/flutter-tflite (Original work published 2023)

Ultralytics. (n.d.). *Home*. Retrieved September 4, 2024, from https://docs.ultralytics.com/

Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace.

Villapol, M. E., Liu, W., Gutierrez, J., Qadir, J., Gordon, S., Tan, J., Chiaraviglio, L., Wu, J., & Zhang, W. (2018). A Sustainable Connectivity Model of the Internet Access Technologies in Rural and Low-Income Areas. In P. H. J. Chong, B.-C. Seet, M. Chai, & S. U. Rehman (Eds.), *Smart Grid and Innovative Frontiers in Telecommunications* (pp. 93–102). Springer International Publishing. https://doi.org/10.1007/978-3-319-94965-9_10

vladimir, Y. (2024). *vladiH/flutter_vision* [Java]. https://github.com/vladiH/flutter_vision (Original work published 2022)

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., Dunnington, D., Brand, T. van den, Posit, & PBC. (2024). *ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics* (Version 3.5.1) [Computer software]. https://cran.r-project.org/web/packages/ggplot2/index.html

# Chapter 4 – Final Remarks

The number of grains within a sorghum panicle is a crucial factor in accelerating breeding programs and estimating sorghum grain yield. This study developed deep learning models and one linear model that can be used to estimate the number of grains from sorghum panicles using images obtained at an on-farm scale. Additionally, the study proposed a mobile application where these models were embedded for real-time usage by farmers, breeders, agronomists, and consultants.

In Chapter 2, a framework linking two Convolutional Neural Networks, and one linear model was developed to estimate the number of grains in a sorghum panicle using images captured by a smartphone camera under field conditions. The framework links two convolutional neural networks and one linear model and represents a significant advancement in the field of agricultural image analysis. The Yolov8 segmentation model demonstrated superior performance and optimal deployment settings when compared to the other tested model. The Sorghum-Net density model exhibited the most effective results among the three tested for grain counting, with no discernible inference time. Lastly, a polynomial linear model was used to link the grain counting from a single side of the panicle to project a number of grains for the entire head.

A multi-operating system mobile application was developed in Chapter 3 for the deployment of the aforementioned models in real-time applications. The developed application comprises two branches: one operating locally in the absence of an internet connection and another functioning remotely when an internet connection is available, and a web server is running. Each branch presents a distinct set of advantages and disadvantages.

It is recommended that future studies endeavor to enhance the three models by incorporating novel (and more diverse) sorghum genotypes (with variation in panicle architecture). With regard to the mobile application, it would be advantageous for future studies to implement a mechanism whereby the application automatically switches between the branches contingent on the availability of an internet connection. Additionally, it would be beneficial to incorporate user settings for historical readings and multiple reading locations for yield estimation.