Zineb Jrondi / Afr.J.Bio.Sc. 6(SI2) (2024) **ISSN:2663-2187**

https://doi.org/ 10.48047/AFJBS.6.Si2.2024.5936-5944

African Journal of Biological Sciences

Journalhomepage: http://www.afjbs.com

Research Paper Open Access

Artificial Intelligence for Agricultural Yield Estimation: A Comprehensive Overview of Recent Approaches

Zineb Jrondi[1], Abdellatif Moussaid[2], Moulay Youssef Hadi[3]

[1] *Laboratory for Computer Science Research, Faculty of Sciences Ibn Tofail University, Kénitra, Morocco. zineb.jrondi@uit.ac.ma*

[2] *ENSIAS, Mohammed V University in Rabat, Rabat, Morocco. abdellatif_moussaid@um5.ac.ma*

[3] *Laboratory for Computer Science Research, Faculty of Sciences Ibn Tofail University, Kénitra, Morocco. Hadi@uit.ac.ma*

Abstract-This paper offers a comprehensive overview of recent advancements in artificial intelligence (AI) algorithms and architectures employed in yield estimation within the agricultural field. The study begins by conducting a thorough science mapping analysis to assess the landscape of yield estimation research, examining trends in published literature to identify pertinent and popular topics. Subsequently, a curated list of popular datasets is presented, accompanied by an elucidation of the lifecycle of utilizing AI for yield estimation. Additionally, the paper undertakes a bibliographic analysis of recent publications, scrutinizing the datasets utilized and the performance scores achieved to provide insights into the state-of-the-art methodologies in this field.

Keywords: Smart Agriculture, Yield Estimation, Artificial Intelligence, Deep Learning, Machine Learning.

1. Introduction

Yield estimation stands as a cornerstone technique for farmers, enabling them to forecast and manage their harvests with precision. By providing accurate insights into expected produce quantities, it empowers farmers to make informed decisions regarding transportation, storage, and distribution, thereby reducing stock shortages and waste while optimizing logistical operations.

Moreover, precise yield estimation grants farmers leverage during sales negotiations. Armed with a thorough understanding of their production capacity, they can strategically plan logistics, secure contracts in advance, and diversify their commercial avenues, ultimately bolstering their economic resilience.

Traditional methods of yield estimation, often reliant on manual observations or surveys, fall short in capturing the spatial variability of crops and are susceptible to subjective biases. This inadequacy can disadvantage farmers during negotiations, leading to less favorable pricing and potential financial losses due to inaccurate estimates.

However, the emergence of artificial intelligence (AI) presents a promising paradigm shift in yield estimation. By harnessing AI techniques alongside remote sensing technologies, farmers can attain more precise and reliable estimates on a finer scale and over broader areas. This convergence of AI and agriculture not only promises enhanced profitability for farmers but also contributes to global food security and agricultural sustainability.

In light of these advancements, this paper endeavors to present a comparative study of the latest approaches in fruit yield estimation over the past four years, focusing on methodologies leveraging fruit detection techniques. Through this exploration, we aim to shed light on the transformative potential of modern technologies in revolutionizing yield estimation practices, paving the way for a more prosperous and sustainable future in agriculture.

In the subsequent sections of this paper, we aim to conduct a science mapping to analyze the evolution of publications in the field of yield estimation. Following this, we will delve into the latest techniques employed for yield estimation, particularly focusing on methods based on fruit detection. Finally, we will draw conclusions based on our findings and insights garnered throughout the study.

2. Science Mapping

Our endeavor to analyze the latest developments in the field of yield estimation led us to use the Connected Papers website. This platform allows for the exploration of scholarly literature by initiating from a highly cited paper and visualizing its interconnectedness with related works. The resulting graph, depicted in Figure 1, offers a comprehensive overview of the landscape of yield estimation research.

Each node in the graph represents a scholarly work, with its size denoting the number of citations it has received. Additionally, the brightness of the green color indicates the recency of the publication; brighter shades signify newer contributions. The interconnections between nodes signify citation relationships, illustrating how research in this domain builds upon prior knowledge and insights [1].

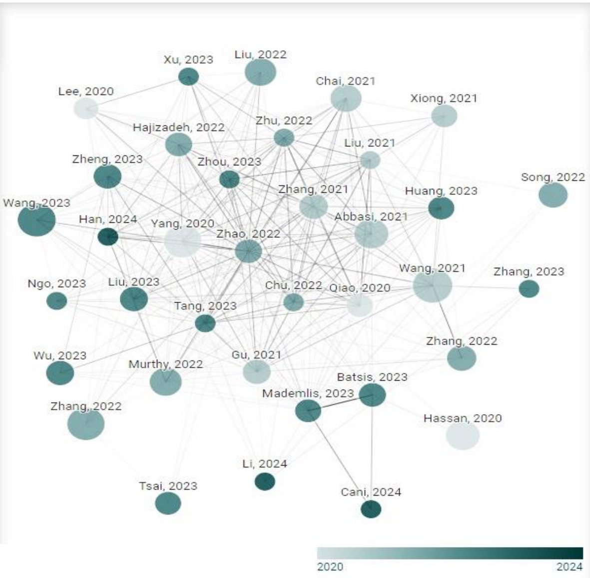

Figure 1 . Science Mapping

Our analysis of the graph spanning from 2020 to 2024 reveals that the field of yield estimation, particularly leveraging artificial intelligence, is witnessing a surge in interest and scholarly activity. This interdisciplinary field has attracted researchers from diverse backgrounds, including computer science, mathematics, and agriculture, reflecting its multifaceted nature and the complexity of the challenges it addresses.

Notably, the graph highlights several prominent nodes characterized by their significant size and vibrant color, indicating their high citation count and recent publication. Among these notable contributions are the works of Wang et al. (2023), Huang et al. (2023), and others. These papers likely represent seminal contributions that have garnered considerable attention within the research community.

Furthermore, our exploration of this science mapping has unearthed a selection of papers that stand out for their influence and relevance. These papers will serve as focal points for our bibliographic study, enabling a deeper examination of the methodologies, findings, and implications within the realm of yield estimation.

In conclusion, the visualization provided by the science mapping offers invaluable insights into the evolution and dynamics of the field of yield estimation. It underscores the emergence of novel concepts and methodologies while signaling areas warranting further exploration and analysis. By leveraging such tools, we aim to contribute to a deeper understanding of this burgeoning field and its potential impact on agricultural practices and technological innovation.

3. Yield Estimation

In the pursuit of implementing yield estimation projects leveraging artificial intelligence, the acquisition of a comprehensive dataset stands as a foundational requirement. Researchers typically adhere to two prominent methodologies, each tailored to suit distinct agricultural contexts (Figure 2). The first approach, widely favored, centers around fruit detection, particularly applicable in arboricultural settings. This methodology necessitates the meticulous collection of field images, subsequently annotated through manual fruit detection processes. Leveraging deep learning object detection techniques, researchers then develop models capable of accurately identifying and delineating fruits within new images, thereby facilitating yield estimation.

Conversely, the second approach pivots towards a holistic collection of parcel-level data, encompassing an array of environmental factors such as climate conditions, fertilization regimes, phytosanitary practices, and irrigation schedules.

Complementing these datasets are spectral insights derived from satellite or drone imagery. Through meticulous labeling of each parcel's yield, researchers employ machine learning algorithms to construct predictive models capable of forecasting yield outputs for individual parcels. This multifaceted approach integrates diverse datasets and analytical techniques, empowering researchers to achieve nuanced and accurate yield estimations tailored to specific agricultural contexts.

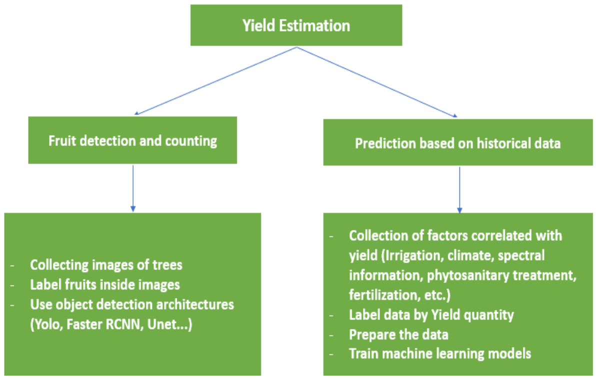

Figure 2. Yield Estimation

Some Popular Dataset

The utilization of Artificial Intelligence (AI) for yield prediction poses inherent challenges in data collection and labeling processes, necessitating substantial time and effort. Collecting comprehensive and accurately labeled datasets is crucial for training AI models effectively. However, this endeavor is often hindered by the complexities of agricultural environments, where factors such as variability in crop growth stages, environmental conditions, and field operations present significant hurdles. Moreover, the manual annotation of large-scale datasets for AI training can be labor-intensive and prone to errors, further exacerbating the challenges associated with data collection [2].

Despite these challenges, Table 2 presents a curated selection of popular datasets meticulously prepared for yield estimation research. These datasets offer researchers valuable resources for testing and validating AI algorithms. By leveraging these meticulously curated datasets, researchers can expedite the development and evaluation of AI models for yield prediction, thereby advancing the state-of-the-art in agricultural technology.

Table 1. Yield Estimation Popular Datasets

Data	species	Source
Orchard Data	Mangoes, Almonds, Apple	http://data.acfr.usyd.edu.au/ag/treecrops/2016-multifruit
Fruit 360	Several fruit	http://www.kaggle.com/moltean/fruits.
Apple dataset images	Apple	https://www.kaggle.com/datasets/basmarg/apple-dataset-images
Corn farming data	Corn	https://www.kaggle.com/datasets/japondo/corn-farming-data

Machine learning and deep learning for yield estimation

The research landscape of yield estimation encompasses a wide array of crop types, each presenting unique challenges and opportunities for accurate prediction. For instance, Table 2 presents the most popular yield estimations works during the last years. In this way, the study [5] focused on estimating orange yield, a crop renowned for its variability in size, shape, and color, posing challenges for traditional yield estimation methods. Leveraging a CNN model, the study achieved an impressive accuracy of 93.8%, underscoring the efficacy of deep learning in capturing the intricate features of citrus fruits.

Similarly, [6] addressed the yield estimation of prunes, a crop characterized by its dense clusters and varying maturity levels. Employing the YOLOv7-Plum architecture, the study attained a remarkable accuracy score of 94.91% using image data. This demonstrates the adaptability of deep learning algorithms in accurately delineating fruit clusters and estimating yield with precision.

Furthermore, [7] explored the estimation of maize yield, a staple crop with complex growth patterns influenced by environmental factors such as temperature, rainfall, and soil moisture. Utilizing machine learning approaches including Cubist, RF, SVM, and XGBoost with satellite data, the study achieved notable results, with Cubist exhibiting the highest performance with an R value of 94.2%. This highlights the potential of remote sensing data in capturing spatial variability and informing predictive models for yield estimation in large-scale agricultural settings.

In terms of data usage, studies employ a diverse range of datasets encompassing both image and numeric data sources. While image data provides valuable visual information for crop identification and delineation, numeric data such as climate variables, soil properties, and agronomic practices offer insights into the underlying factors influencing yield. For instance, [10] utilized a CNN-LSTM model for soybean yield estimation, leveraging their own data containing both image and numeric variables. This integration of multi-modal data sources enables the model to capture both spatial and temporal dependencies, enhancing prediction accuracy.

Moreover, [12] employed the Faster R-CNN architecture for coconut yield estimation, utilizing image data to identify and segment coconut palms in aerial imagery. This approach enables precise delineation of individual coconut palms and estimation of yield based on canopy size and density.

In summary, the diversity of crops, coupled with the utilization of various data types and sophisticated deep learning architectures, underscores the multidimensional nature of yield estimation research. By leveraging advanced AI techniques and comprehensive datasets, researchers can address the complexities of agricultural systems and contribute to enhancing crop productivity and food security on a global scale.

Table 2. Yield Estimation Benchmark

Ref	***Crop***	***Approach***	***Score***	***Data type***
[4]	Apple	Mask R-CNN	73% (Accuracy)	Images data
[5]	Orange	CNN	93,8% (Accuracy)	Images data
[6]	Prunes	YoloV7-Plum	0.94 (Accuracy)	Images data
[7]	Maize	Xgboost	0.85(R-square)	Satellite data
[8]	wheat	Regression model	0.228 (RMSE)	Numeric data
[9]	soybean	CNN-LSTM model	329,53 (RMSE)	Images data
[10]	Mangoes	YOLOv2	0.96(Precision)	Images data
[11]	Tomato	Mask R-CNN	0.88(Precision)	Images data
[12]	Coconuts	Faster R-CNN	0.89(Precision)	Images data
[13]	Wheat ears	YOLOv4	0.98(R-square)	Images data
[14]	Wheat ears	Retina Net	0.92(Accuracy)	Images data
[15]	Tomato	YOLOv5	0.85 (Precision)	Images data
[16]	Apple	YOLOv4	0.91(Accuracy)	Images data
[17]	Cotton seedling	CenterNet	0.982 (F1-score)	Images data
[18]	Rapeseed	YOLOv5	0.96(R-square)	Images data
[19]	Tomato	YOLOv5	0.97(Precision)	Images data
[20]	Lettuce	YOLOv5	0.98(Precision)	Images data

4. Conclusion

In conclusion, our study embarked on a comprehensive exploration of yield estimation methodologies within the agricultural domain, with a focus on leveraging artificial intelligence (AI) techniques. Through the use of the science mapping technique, we obtained a global idea about the landscape of yield estimation research, highlighting its significance in enhancing precision farming practices. The visualization of scholarly literature and citation relationships underscored the interdisciplinary nature of this field and the growing interest among researchers from diverse backgrounds.

Our analysis revealed the pivotal role of AI algorithms in revolutionizing yield estimation, offering farmers unprecedented accuracy and insights into their production capacities. By harnessing deep learning and machine learning techniques, researchers have achieved remarkable results in accurately predicting yields across various crop types. However, it is essential to acknowledge the challenges inherent in this endeavor, particularly regarding data collection and model generalization.

While our benchmarking showcased impressive performance scores across different methodologies and crop types, it is crucial to recognize that achieving such results requires meticulous data annotation, model training, and validation processes. Moreover, the scalability of these methodologies to different agricultural contexts remains a significant consideration. The complexities of agricultural environments, coupled with variations in crop types, growth patterns, and environmental conditions, necessitate tailored approaches and continuous refinement of models. It is evident that the transformative potential of AI in yield estimation holds promise for enhancing agricultural productivity, optimizing resource utilization, and mitigating risks associated with crop management. However, realizing this potential requires collaborative efforts from researchers, policymakers, and agricultural stakeholders to address existing challenges and ensure the widespread adoption of AI-driven solutions in farming practices.

References

[1] A. Moussaid, S. E. Fkihi, and Y. Zennayi, “Citrus Orchards Monitoring based on Remote Sensing and Artificial Intelligence Techniques: A Review of the Literature:,” in Proceedings of the 2nd International Conference on Advanced Technologies for Humanity, Rabat, Morocco: SCITEPRESS - Science and Technology Publications, 2020, 172–178. 10.5220/0010432001720178.

[2] C. C. Ukwuoma, Q. Zhiguang, M. B. Bin Heyat, L. Ali, Z. Almaspoor, and H. N. Monday, “Recent Advancements in Fruit Detection and Classification Using Deep Learning

Techniques," Math. Probl. Eng., 2022, 1–29, 2022, 10.1155/2022/9210947.

[3] Y. Huang, Y. Qian, H. Wei, Y. Lu, B. Ling, and Y. Qin, "A survey of deep learning-based object detection methods in crop counting," Comput. Electron. Agric., 215, 108425, 2023, doi: 10.1016/j.compag.2023.108425.

[4] M. Ferrer-Ferrer, J. Ruiz-Hidalgo, E. Gregorio, V. Vilaplana, J.-R. Morros, and J. Gené-Mola, "Simultaneous fruit detection and size estimation using multitask deep neural networks," Biosyst. Eng., vol. 233, pp. 63–75, Sep. 2023, doi: 10.1016/j.biosystemseng.2023.07.010

[5] S. Zeeshan, T. Aized, and F. Riaz, "The Design and Evaluation of an Orange-Fruit Detection Model in a Dynamic Environment Using a Convolutional Neural Network," Sustainability, vol. 15, no. 5, p. 4329, Feb. 2023, doi: 10.3390/su15054329.

[6] R. Tang, Y. Lei, B. Luo, J. Zhang, and J. Mu, "YOLOv7-Plum: Advancing Plum Fruit Detection in Natural Environments with Deep Learning," Plants, vol. 12, no. 15, p. 2883, Aug. 2023, doi: 10.3390/plants12152883.

[7] X. Chen, L. Feng, R. Yao, X. Wu, J. Sun, and W. Gong, "Prediction of Maize Yield at the City Level in China Using Multi-Source Data," Remote Sens., vol. 13, no. 1, p. 146, Jan. 2021, doi: 10.3390/rs13010146.

[8] D. Ashourloo, M. Manafifard, M. Behifar, and M. Kohandel, "Wheat Yield Prediction based on Sentinel-2, Regression and Machine Learning Models in Hamedan, Iran," Sci. Iran., vol. 0, no. 0, pp. 0–0, Jun. 2022, doi: 10.24200/sci.2022.57809.5429.

[9] J. Sun, L. Di, Z. Sun, Y. Shen, and Z. Lai, "County-Level Soybean Yield Prediction Using Deep CNN-LSTM Model," Sensors, vol. 19, no. 20, p. 4363, Oct. 2019, doi: 10.3390/s19204363.

[10] J. Xiong et al., "Visual detection of green mangoes by an unmanned aerial vehicle in orchards based on a deep learning method," Biosyst. Eng., vol. 194, pp. 261–272, Jun. 2020, doi: 10.1016/j.biosystemseng.2020.04.006

[11] "Tomato_Fruit_Detection_and_Counting_in_Greenhouses.pdf.

[12] S. Parvathi and S. Tamil Selvi, "Detection of maturity stages of coconuts in complex background using Faster R-CNN model," Biosyst. Eng., vol. 202, pp. 119–132, Feb. 2021, doi: 10.1016/j.biosystemseng.2020.12.002

https://doi.org/ 10.48047/AFJBS.6.Si2.2024.5936-5944

[13] B. Yang, Z. Gao, Y. Gao, and Y. Zhu, "Rapid Detection and Counting of Wheat Ears in the Field Using YOLOv4 with Attention Module," Agronomy, vol. 11, no. 6, p. 1202, Jun. 2021, doi: 10.3390/agronomy11061202.

[14] C. Wen et al., "Wheat Spike Detection and Counting in the Field Based on SpikeRetinaNet," Front. Plant Sci., vol. 13, p. 821717, Mar. 2022, doi: 10.3389/fpls.2022.821717.

[15] Y. Egi, M. Hajyzadeh, and E. Eyceyurt, "Drone-Computer Communication Based Tomato Generative Organ Counting Model Using YOLO V5 and Deep-Sort," Agriculture, vol. 12, no. 9, p. 1290, Aug. 2022, doi: 10.3390/agriculture12091290.

[16] F. Gao et al., "A novel apple fruit detection and counting methodology based on deep learning and trunk tracking in modern orchard," Comput. Electron. Agric., vol. 197, p. 107000, Jun. 2022, doi: 10.1016/j.compag.2022.107000.

[17] H. Yang et al., "Multi-object tracking using Deep SORT and modified CenterNet in cotton seedling counting," Comput. Electron. Agric., vol. 202, p. 107339, Nov. 2022, doi: 10.1016/j.compag.2022.107339.

[18] J. Li et al., "Automatic counting of rapeseed inflorescences using deep learning method and UAV RGB imagery," Front. Plant Sci., vol. 14, p. 1101143, Jan. 2023, doi: 10.3389/fpls.2023.1101143.

[19] J. Rong, H. Zhou, F. Zhang, T. Yuan, and P. Wang, "Tomato cluster detection and counting using improved YOLOv5 based on RGB-D fusion," Comput. Electron. Agric., vol. 207, p. 107741, Apr. 2023, doi: 10.1016/j.compag.2023.107741.

[20] P. Zhang and D. Li, "Automatic counting of lettuce using an improved YOLOv5s with multiple lightweight strategies," Expert Syst. Appl., vol. 226, p. 120220, Sep. 2023, doi: 10.1016/j.eswa.2023.120220.