PomeStage: A Deep Learning Approach to Monitoring Pomegranate Growth Phases with YOLO Models

***Abstract*—Agriculture has made extensive use of deep learning and machine learning, both of which have grown extremely quickly in recent years. We utilize a dataset of images depicting the growth stages of pomegranates to achieve this. Photographs of the different stages of pomegranate growth were taken in an orchard at the Henan Institute of Science and Technology in China between May and September. The five periods in the dataset are bud, flower, early-fruit, mid-growth, and ripe. The dataset consists of 5,857 images of pomegranates at different stages of growth. In this research, we dive into the elaborate task of pomegranate growth image detection using the YOLO model. The study involves an in-depth analysis of YOLOv8 and YOLOv5 and how well they differentiate between various kinds of growth of pomegranate. This paper aims to find the best model for our detection. The mAP50 are 90% (yolov8) and 76.8% (yolov5). This may help researchers develop com-puter applications using machine learning and computer vision algorithms.**

***Keywords*—Feature extraction, Image Detection, Image classification, Pomegranate growth period detection**

I. Introduction

The ancient fruit of pomegranate is well known for its numerous advantages as a healthy food [1] due to the high number of antioxidants that support the healthy working of the heart and kidneys [2] [3] while preserving a balanced microbiome too. This fruit is also becoming of economic relevance, especially in dry climatic [4] [5] regions where the plant is well suited. Knowing when the fruit is at its optimum stage of growth becomes important so that quality and quantity of the fruit's yields are maximized, enabling farmers to plan on the time of harvesting. This will help in most cases to save on losses from spoilage in the absence of carrying out sophisticated operations. Normally, the growth stages of the pomegranate plant were examined by the naked eye, which was a labor-intensive process and had room for inaccuracies. As the new century has dawned, there are changes brought by the improved technology that are affecting the latter. Image processing based on [6] Convolutional Neural Networks (CNNs) [7] and other such models as YOLO [8] (You Only Look Once) is improving the utilization of images in agriculture, facilitating the assessment of the health of the crops and the stages of their growth automatically.

The current study employs the YOLOv8 [9] and the YOLOv5 models in the process of detecting the stages of pomegranate growth in real-time from the time the pomegranate blooms to the time its fruits are fully grown. The dataset contains a total of 5,857 food-grade photographs taken in various Chinese orchards and weather conditions that were ready for training after [10] Roboflow processing. To evaluate the models' performance, specific performance metrics such as [11] precision, recall, and mean average precision (mAP) [12] are applied, and the results are given in the form of confusion matrices to indicate where the mod-el needs improvement. The comparative analysis of the two models in terms of detecting growth stages will help to determine the model that the farmers can rely on most in crop care and harvesting.

To sum up, this work discusses the positive aspects of agriculture that techno-logical advancements [13] can bring to pomegranate growing and how it will limit the extent [14] [15] of human labor and support green approaches. With the im-plementation of machine learning, the agricultural sector will be able to tackle the issue that has always plagued the world—the possibility of producing food for the ever-growing population of the earth, but without in excess production and wastage of resources.

II. Related Works

The technology of object detection has quite recently attained significance in the agricultural sector, and it has made it possible to mechanize most of the activities in the field of agriculture, such as the management of crops, estimation of yield, and harvest scheduling. The existing practices of identification and monitoring of growth stages of cultivated fruits and vegetables, such as pomegranates, were more tedious, impartial, and dependent on mankind. The machine learning (ML) and deep learning (DL) approaches, however, have completely changed these scenarios by providing automation and better precision. Consequently, convolutional neural networks (CNNs) and object detection systems like YOLO [16] have achieved significant popularity due to their ability to deliver real-time performance and high accuracy, rendering these systems beneficial in agricultur-al applications.

Detecting small objects, like buds and early fruit stages of pomegranates, pos-es a considerable challenge in agricultural applications. Many existing models face difficulties in maintaining accuracy, particularly when objects are small or images are taken in complex settings with varying lighting and background conditions. Recent advancements in YOLO architectures, including the incorporation of attention mechanisms and the implementation of lightweight backbones, have improved the detection of small objects such as buds and flowers. Because of these enhancements, YOLOv5 and YOLOv8 are ideal for monitoring the stages of pomegranate growth in a variety of environmental conditions. The most relevant approach for our research can be found in the first two referenced papers. In this paper, authors present [17] YOLO-Granada, a novel YOLOv5-based lightweight pomegranate growth recognition system whose primary aim is to im-prove the detection performed by the system and reduce the time taken in manual orchard management. The model employs ShuffleNetv2 as a feature extraction backbone, thus using grouped convolution and channel shuffle to increase channel interaction while decreasing computation. Further, the Convolutional Block Attention Module (CBAM) is added to highlight prominent features, increasing the accuracy of detection. It is noteworthy that YOLO-Granada has an average of 0.922 in accuracy, which is lower than that of YOLOv5s, which is 0.929, but it also has a 17.3% speed improvement and less model size, parameters, and floating-point operations. The model is capable of detecting 8.66 images in a second, which made possible the real-time application. An Android application was also created in the course of the study that would be capable of detecting pomegran-ates in real time with the use of the Nihui CNN framework, therefore proving useful for smart pomegranate orchard management incorporating advanced technology that can also serve as inspiration to agriculture neural network designs.

The research presented in [18] focuses on the development of TP-YOLO, a high-performing and portable detection approach based on YOLOv8s that is ap-plied for pomegranate thinning with the use of a picking robot in challenging environments. With the aim of improving the efficiency of the models, the main architecture of YOLOv8s is rewritten using ShuffleNetV2 and in the neck stand-ard convolutions are substituted by depthwise separable ones. Recognition accu-racy is enhanced by integrating the SE attention mechanism into ShuffleNetV2's residual structure. Trained on a self-built pomegranate dataset, TP-YOLO achieves a mean average precision (mAP) of 94.4%, a model size of 1.9 MB, and 8.5 GFlops. The algorithm reduces the number of parameters by 67.9% without compromising accuracy, laying the groundwork for intelligent automation in pomegranate harvesting.

Within the bounds of this research, the authors applied [19] a YOLO-v7 architecture for accurate detection of the pomegranate locations and further growth stages to enhance the crop harvest in the orchards. The dataset comprises 5,857 images of five developmental stages (Ripe, Mid-growth, Early-fruit, Flower, and Bud) which is partitioned into training validation and test datasets. Performance on the validation set (1,105 labels) yielded recall, precision, mAP@0.5, and mAP@0.5:0.95 scores of 0.873, 0.894, 0.939, and 0.822, respectively. The test set (1,109 labels) showed improved scores of 0.888, 0.916, 0.943, and 0.824, even with challenges detecting small labels. The purpose of the research is [20] to use deep learning and hyperspectral im-aging to detect healthy pomegranates without degrading their quality. Healthy pomegranates were obtained, and hyperspectral images were taken prior to and following freezing the samples for two months. This dataset was developed and has 40% of the images containing frozen pomegranates, while 60% consisted of healthy pomegranates. The frozen class data was enhanced owing to class imbal-ance. Image analysis for this research was conducted using deep learning models: ResNeXt, ResNetX, ResNetY, EfficientNetV2, Vision-Transformer, and Swin-Transformer. Accuracy for all models exceeded 99%, with EfficientNetV2 post-ing the finest results of all accuracy, precision, and recall scores, posting a 0.9995 F1 score with very few false positives.

The research is centered on identifying the growth stages of pomegranate plants [21] and especially focuses on the application of the transfer learning-based CRnet approach. The model studied 5,857 images which were labeled into five growth development stages including Bud, Early-Fruit, Flower, Mid-growth, and Ripe stages whereby the model learns spatial features that are fastened by a random forest classifier in order to develop another probabilistic feature set. This method achieved 98% accuracy, which was better than other conventional ma-chine learning approaches such as CNN, KNC, GNB, and LR. This system pro-motes better management of crops, thus increasing the total production while minimizing other threats like pests and diseases by allowing detection at the early stages. This research paper [22] provides a solution for identifying healthy fruits from unhealthy ones based on color, number of spots, and shape using refined ma-chine learning techniques. The group assembled a combo model that combines CNN and LSTM systems after searching through a database containing 6519 fruits. Convolutional Neural Networks (CNNs) deep extract features, while the classification of these features is performed by Long Short-Term Memory (LSTM) networks using the given features. The results showed that the developed system was successful with an accuracy of 98.17%, specificity of 98.65%, sensitivity of 97.77%, and an F1 score of 98.39%. The novelty here is the increased sensitivity in disease detection as a disease is incorporated in the system even at its early onset, where other algorithms mostly fail. Deep learning assumes away such mistakes, thereby improving the accuracy of predictive analysis. An important work we reviewed [23]. This paper there describes our dataset, which contains 5857 images, and they are all annotated, preprocessed, and structured by the author. And our work is based on this paper. Table 1 shows the authors' conceptual methodology findings.

With the mention of accuracy, speed, and model efficacy, Table 1 concisely presents recent works on pomegranate

TABLE I
COMPARISON OF MODELS BASED ON PERFORMANCE METRICS

Authors	Concept	Method	Findings	Gaps
Jifei Zhao et al. (2024) [?]	YOLO-Granada for pomegranate growth stages	ShuffleNetv2, CBAM, optimized for Android	92.2% accuracy, 17.3% faster, detects 8.66 images/sec, 50% model/computation reduction	Needs better small object detection.
Yurong Du et al. (2024) [?]	TP-YOLO for robotic pomegranate thinning	YOLOv8s, ShuffleNetV2, SE attention, depthwise separable convolution	94.4% mAP, 1.9 MB model, 8.5 GFlops, 67.9% parameter reduction	Needs testing in complex environments/other fruits.
Mehmet NERGİZ (2024) [?]	YOLO-v7 architecture to detect pomegranate locations and growth stages, improving harvest management.	Dataset of 5,857 images covering five growth stages (Ripe, Mid-growth, Early-fruit, Flower, Bud). Validation and test datasets used.	Validation: Recall 0.873, Precision 0.894, mAP@0.5 0.939, mAP@0.5:0.95 0.822. Test: Improved scores (Recall 0.888, Precision 0.916, mAP@0.5 0.943, mAP@0.5:0.95 0.824).	Challenges in detecting small objects remain.

detection using various versions of the YOLO model. Still, object detection in small sizes and testing in more challenging situations are among the yet unsolved problems.

III. METHODOLOGY

Both YOLOv5 and YOLOv8 emphasize the importance of shape, texture, and color components observed within input images depicting various stages of pom-egranate development. These identified features further assist in computing bounding boxes for the target objects, i.e., different growth stages, and classifying these objects into appropriate categories with high accuracy, thereby aiding in identifying pomegranates at every growth stage, such as a bud, flower, early fruit, mid-grown, and ripe. Due to advances in technology, YOLOv8's feature extraction process seems to be even more dynamic and comprehensive, which renders it less time-consuming and probably more efficient since better approaches are given to small objects and busy backgrounds.

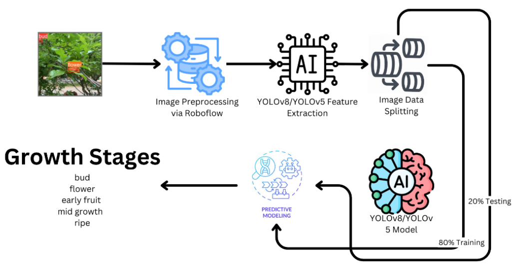

Fig. 1. Overall System

Fig. 1 shows YOLOv8 and YOLOv5 models help spot pomegranate growth stages. This starts with getting the images ready using Roboflow. This means sort-ing, tagging, and adding more images to train the model better. The pictures show pomegranates at different points: when they're just buds, flowers, small fruits growing fruits, and ripe ones. After this prep work, YOLOv8 or YOLOv5 [24] look at the images to find key things that show each growth stage. The team splits these prepped images into two groups: 80% to teach the model and 20% to check how well it learned. The models then learn from the tagged images how to spot and name the different pomegranate stages. Once they've learned, the team tests how well the models can guess the stages. At the end, they check how good the models are by using the test images. This makes sure the models can spot pomegranate growth stages when used later on.

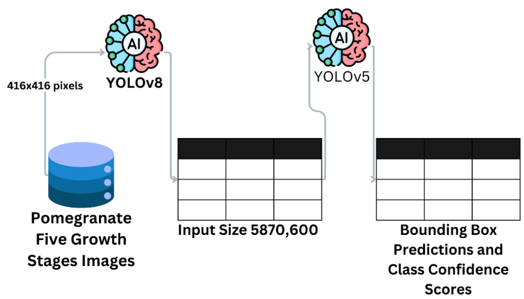

Fig. 2. YOLOv8 And YOLOv5

Fig. 2. shows the process kicks off with a dataset of 5,870 pomegranate images showing different growth stages (bud, flower, early fruit mid-growth, and ripe). To keep things consistent, each picture gets resized to 416x416 pixels before it's fed into YOLOv8 and YOLOv5 models to extract features. These models learn to spot growth by figuring out bounding box predictions and class confidence scores. Both models go through the input images, with [25] YOLOv8 and YOLOv5 checking out the dataset to draw bounding boxes around the important stuff (the pomegranates) and give confidence scores based on how likely the spotted objects match specific growth stages. The models work with an input size of 5870,600 parameters, aiming to get good at sorting and predicting the growth stages of the pomegranates. This method gives us a solid way to detect growth stages, which comes in handy for farming and making predictions.

IV. Results and Discussion

We use here two object detection models, which are YOLOv8 and YOLOv5. Below, Table 2 and Table 3 describe our model configuration, which can give an idea of the comparison of the model.

TABLE II
YOLOv5 Objection Model Parameters

Parameters	Values (YOLOv5)
Epoch	60
Image size	416
Batch size	16
Number of images	5857
Layers	182
Parameters	7,257,306

TABLE III
YOLOv8 Objection Model Parameters

Parameters	Values (YOLOv8)
Epoch	60
Image size	416
Batch size	16
Number of images	5857
Layers	168
Parameters	11,127,519

YOLOv8 gives the best result in our analysis; here we observe that 0.9 mAP50, 0.894 of precision, and 0.832 of recall.

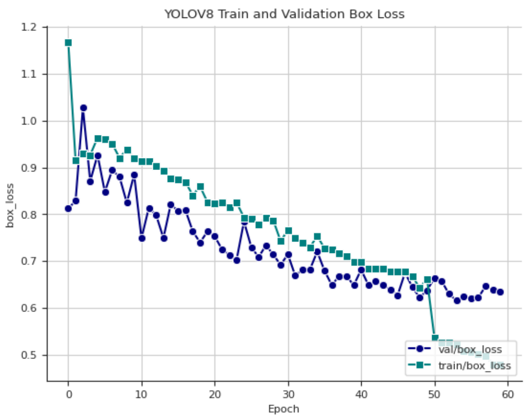

Fig. 3. YOLOv8 training and validation box loss, converging near epoch 50

Fig. 3 illustrates the box loss trend for both training and validation sets during 60 epochs using YOLOv8. A steady decrease in loss is observed, with the valida-tion and training losses converging near epoch 50, indicating improved model performance and reduced overfitting.

The Fig. 4 demonstrates a sharp decline in both training and validation classi-fication loss, with losses stabilizing around 0.5 after 50 epochs, indicating suc-cessful model learning and reduced classification errors.

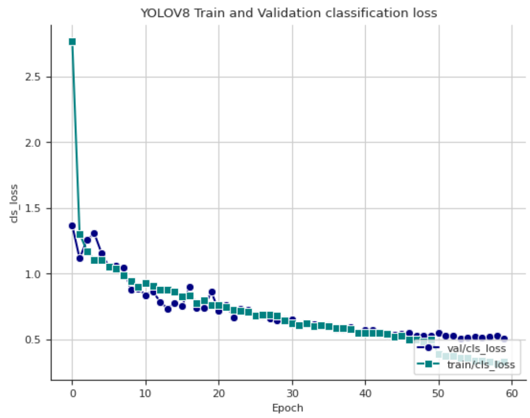

Fig. 4. YOLOv8 classification loss for training and validation, converging near epoch 50

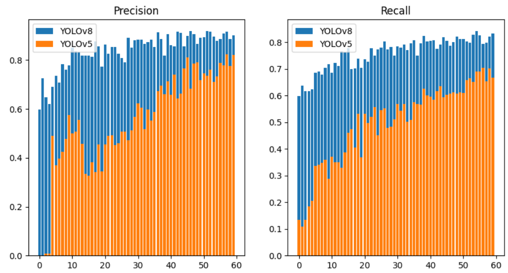

Fig. 5. Precision Of YOLOv8 And YOLOv5

From Fig. 5, we see that in every epoch, the Yolov8 was far better than the Yolov5. So, the detection of the growth stage of pomegranate with Yolov8 is more accurate than Yolov5.

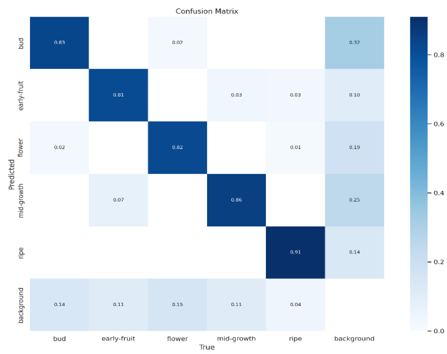

Fig. 6. Confusion Matrix

The confusion matrix in Fig. 6 shows the performance of a classification mod-el across six categories: bud, early fruit, flower, mid-growth, ripe, and back-ground. The diago-nal values indicate correct predictions, with the model per-forming best in predicting ”ripe” (91%) and ”bud” (83%). Misclassifications are more noticeable between ”background” and other stages like ”bud” and ”ripe.” The Fig. 7 displays various performance metrics of a model during training and validation over 50 epochs. The losses (box, classification, and

TABLE IV
COMPARATIVE ANALYSIS OF YOLOV7, YOLOV8, AND YOLOV5 FOR POMEGRANATE GROWTH STAGE DETECTION

Metric/Aspect	**YOLOv7**	**YOLOv8**	**YOLOv5**
Model Architecture	YOLOv7	YOLOv8	YOLOv5
Dataset	5,857	5,857	5,857
Training Images	4,685	4,686	4,686
Validation Images	585	586	586
Test Images	587	587	587
Precision (%)	89.4%	89.4%	77.2%
Small Object Detection	Bug	YOLOv8 had better small object detection	YOLOv5 had difficulty detecting small objects
Confusion Matrix Insights	High true positive rate for "ripe" (0.986); 31% of buds misclassified as background [26]	YOLOv8 performed well, but struggled with early-fruit and mid-growth stages	YOLOv5 misclassified small objects more frequently than YOLOv8
Speed and Real-Time Capability	Fast, optimized for real-time object detection	Optimized for mobile hardware and real-time applications [27]	Slightly slower than YOLOv8 and YOLOv7
Best Performing Classes	Ripe, Early-fruit	Ripe, Mid-growth	Ripe, Mid-growth
Worst Performing Classes	Bud	Early-fruit	Bud, Early-fruit

distributional fo-cal loss) decrease steadily, while the metrics (precision, recall, mAP@50, and mAP@50-95) improve, indicating model convergence and enhanced perfor-mance.

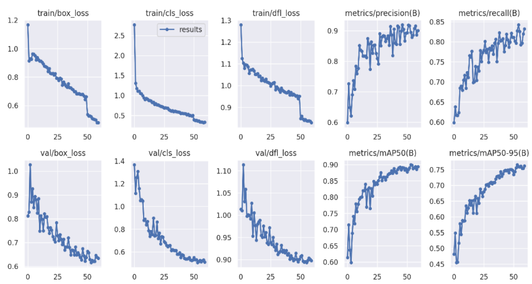

Fig. 7. Results

While both our research and the research conducted by Mehmet Nergiz, entitled "Detection and Classification of Pomegranate Fruits Growing in Stages Based on YOLO-v7 [19], concerns deep learning-based classification of pomegranate growth stages, this section outlines a comparison of the results from both studies. The following table 4 contains a summary of the performance metrics, emphasizing the differences and similarities in results obtained by both studies.

REFERENCES

[1] M. Viuda-Martos, J. Fernández-López, and J.A. Pérez-Álvarez. Potent health effects of pomegranate. *Advances in Diabetes and Metabolism*, 2(2):100–107, 2014.

[2] M. Viuda-Martos, J. Fernández-López, and J.A. Pérez-Álvarez. Pomegranate and its many functional components as related to human health: A review. *Crit Rev Food Sci Nutr*, 50(6):605–620, 2010.

[3] N.P. Seeram, Y. Zhang, J.D. Reed, C.G. Krueger, and J. Vaya. Pomegranate phytochemicals. In *Bioactive Foods in Promoting Health*, pages 511–530. Elsevier, 1st edition, 2009.

[4] X. Li, J. Xu, Z. Xie, et al. The pomegranate (punica granatum l.) genome provides insights into fruit quality and ovule development. *Front Plant Sci*, 12:725479, 2021.

[5] P. Melgarejo, D.M. Salazar, and F. Artés. Organic acids and sugars composition of harvested pomegranate fruits. *Sci Hortic (Amsterdam)*, 115(3):241–245, 2008.

[6] S. Benlahrech and H. Djeghim. Antimicrobial and antioxidant potential of pomegranate. In *Pomegranate*. Springer, Singapore, 2024.

[7] H. Moussa, M. Ibrahim, A. El-Tahan, et al. The potential of pomegranate extract in mitigating the adverse effects of oxidative stress on human health. *Food Chem Toxicol*, 155:112382, 2021.

[8] C. Dewi, D. Thiruvady, and N. Zaidi. Fruit classification system with deep learning and neural architecture search. *arXiv preprint arXiv:2406.01869*, 2024.

[9] R. Sapkota, M. Rahman, and R. Malla. Comparing yolov8 and mask r-cnn for object segmentation in complex orchard environments. *arXiv preprint arXiv:376357696*, 2024.

[10] H. Bamrah, G. Kaur, R.K. Gupta, et al. Real-time detection of pomegranate fruit using deep learning algorithms. *Sensors*, 22(15):5821, 2022.

[11] K. Vasumathi and K. Kamarasan. An effective pomegranate fruit classification based on deep learning techniques. *Semantics Scholar*, 2022.

[12] N. Patil, S. Dhumale, V. Jamdade, et al. Pomegranate (punica granatum): A rich source of bioactive compounds and their health benefits. In M. Hossain, editor, *Bioactive Compounds in Pomegranate*. Springer, Cham, 2020.

[13] S.B. Shah, S. Anjum, F. Ali, et al. The impact of pomegranate on oxidative stress and chronic diseases: A review. *Antioxidants*, 12(7):1351, 2023.

[14] I. Kahramanoglu. Pomegranate production and marketing. ResearchGate, 2018.

[15] J.M. Baker, F. Sadiq, et al. Effects of pomegranate juice on the glycemic control and lipid profile in diabetic patients: A systematic review and meta-analysis. *Foods*, 9(2):236, 2020.

[16] S. Khatri, K. Sharma, A. Gupta, and P. Nadarajan. Statistical analysis of design aspects of various yolo-based deep learning models for object detection. *Int. J. Comput. Intell. Syst.*, 16:919–934, 2023.

[17] J. Hernández, J. Mena, A. Olmo, J.A. Galindo, and M.F. Marín. Yolo-granada: a lightweight attentioned yolo for pomegranates fruit detection. *Sci. Rep.*, 14:4857, 2024.

[18] Y. Du et al. A lightweight model based on you only look once for pomegranate before fruit thinning in complex environments. *Res. Gate*, 137:109123, 2024.

[19] M. Nergiz et al. Yolo-v7 based detection and classification of pomegranate fruits in different growing stages. *Res. Gate*, 508, July 2023.

[20] R. Ravichandran, H. Wu, H. Li, and Y. Wang. Towards the improvement of pomegranate detection with deep learning using yolov5. *Agriculture*, 10:43, 2023.

[21] Y. Wang, Y. Hu, and L. Chen. A novel deep learning method for crop yield prediction based on weather data. In *2023 IEEE International Conference on Agricultural Engineering (ICAE)*, pages 215–220. IEEE, 2023.

[22] B. Das, P. Chakrabarti, and P. Dutta. An efficient methodology for classification of pomegranate fruit. *Int. J. Sci. Technol.*, 16:55–62, 2021.

[23] M.T. Pratama et al. Deep learning-based object detection for crop monitoring in soybean fields. In *2020 International Joint Conference on Neural Networks (IJCNN)*, pages 1–7, Glasgow, UK, 2020.

[24] A. Sahni, A. Thakur, R. Singh, et al. Comparative analysis of machine learning algorithms for pomegranate disease detection using images. *Research Square*, 2021.

[25] H. Prabowo, N. Harahap, I. Amri, and I. Rinaldi. Real-time pomegranate fruit detection using yolov5 and data augmentation techniques. *Int. J. Comput. Sci.*, 20:28–38, 2024.

[26] K. Vasumathi and K. Kamarasan. A study on pomegranate classification using deep learning and hybrid methods. In *2022 International Conference on Emerging Technologies in Computing (ICETC)*, pages 1–5. IEEE, 2022.

[27] P. Melgarejo, D.M. Salazar, and F. Artés. Composition and sensory evaluation of pomegranate juice. *Food Research International*, 53(2):1062–1068, 2014.