Strawberry Classifier

Mariah Muscat
Institute of Information & Communication Technology
Malta College or Arts, Science & Technology
Corradino Hill
Paola PLA 9032
Mariah.Muscat.D57255@mcast.edu.mt

***Abstract*—**

***Index Terms*—**

I. Introduction

The vibrant red strawberry isn't merely a seasonal delight. In Malta, it has a distinct growing season, flourishing during the late winter and spring months. This unique window makes the arrival of Maltese strawberries a highly anticipated event for both residents and tourists. The popularity is further underscored by the success of the annual strawberry festival, a celebration where this delectable fruit takes center stage. From enjoying them fresh off the vine to incorporating them into innovative culinary creations, the festival showcases the versatility of the strawberry.

However, the journey from a humble seed to a juicy, edible berry is far from simple. Cultivating strawberries requires meticulous care and expertise. Unlike other fruits, they are delicate and susceptible to over-watering. Even seemingly beneficial rain can become detrimental if it exceeds the plant's needs [1]. This inherent fragility underscores the dedication and skill required by Maltese strawberry farmers.

The rise of technology has permeated every aspect of our lives, and agriculture is no exception. This study aims to leverage advancements in this field to address a crucial challenge in Maltese strawberry production identifying the optimal harvesting time. By implementing a machine learning-based solution to analyze various datasets, we propose a system that can accurately determine when a strawberry has reached its peak ripeness and is ready to be enjoyed by consumers.

Grounded in a pragmatic philosophy, this study emphasizes the practical application of research findings to address real-world challenges in strawberry production [2]. Employing a deductive approach, the study begins with a theoretical framework derived from existing literature and extends it to develop practical solutions for strawberry farming. The research strategy [2] incorporates both quantitative and qualitative methods, with quantitative analysis used for dataset construction and algorithm evaluation, while qualitative insights from experienced growers contribute to refining the system's recommendations, ensuring their anonymity. Various methodologies are employed, including data collection through image processing techniques and algorithm development using machine learning approaches. Operating within a limited time frame, the study focuses on the immediate application of findings to enhance strawberry production in Malta. Central to the study are cutting-edge techniques such as computer vision and machine learning, enabling the development of a sophisticated smart strawberry system [3] .

By traversing through the layers of the research onion, this study aims to provide a comprehensive understanding of strawberry growth stages and interventions, ultimately contributing to the advancement of agricultural technology and enhancing the sustainability of strawberry farming practices in Malta

This research proposes a "smart strawberry system" powered by computer vision. By analyzing digital details within strawberry images, the system goes beyond appearance to predict the ideal harvest window, minimizing waste and maximizing yield [3]. This early detection system also identifies potential diseases lurking beneath the surface, allowing farmers to take preventative measures and ensure healthy strawberries. Similar to how image processing aids pest management, this technology offers earlier disease detection, benefiting both farmers and consumers. In essence, the study aims to develop a system that detects strawberry growth stages and diseases using computer vision, offering recommendations based on these findings.

This study explores into the exploration of artificial intelligence (AI) techniques' viability in categorizing strawberry growth stages and proposing interventions to enhance growth. It hypothesizes that an AI-driven system can be fashioned to streamline this process, thereby enhancing the precision and efficiency of strawberry production. Specifically, the study endeavors to construct and assess a system proficient in automatically discerning diverse strawberry growth stages utilizing visual data and subsequently suggesting suitable interventions, including watering, nutrient application, and temperature adjustments, tailored to the identified growth stage. To fulfill these aims, the research questions are the following:

1) How can a robust and comprehensive dataset be constructed that accurately captures the visual characteristics of different strawberry maturity stages?
2) What AI algorithm is most effective for classifying the various growth stages of a strawberry plant?
3) How do the system's recommendations for growth interventions compare to the advice provided by experienced strawberry growers or established best practices?

II. Literature

A. YOLO & Roboflow: Empowering Agricultural Object Detection

YOLO (You Only Look Once) object detection models have emerged as a powerful tool for real-time computer vision tasks in agriculture due to their speed and accuracy. However, training custom YOLO models often presents a significant challenge, requiring expertise in data annotation and management. Roboflow, a cloud-based platform, addresses this bottleneck by offering user-friendly tools for data annotation, augmentation [4], and integration with YOLO frameworks.

Studies by Singh et al., 2023 [5] and Li et al., 2017 [6] showcase YOLO's effectiveness in agricultural object detection. Singh et al. achieved a promising mAP (mean Average Precision) of 51% for fruit and vegetable recognition in a real-world market setting, while Li et al. demonstrated a 90.42% accuracy for apple leaf disease detection using a simulated dataset. These results highlight YOLO's potential for real-time tasks like crop health monitoring, yield estimation, and livestock management.

Roboflow, [7] simplifies the deep learning development lifecycle, including data versioning, experiment tracking, and model deployment. This can significantly reduce the time and effort required to train YOLO models for various agricultural applications. By combining YOLO's real-time performance with Roboflow's user-friendly approach, researchers and practitioners can develop and deploy object detection models more efficiently, leading to improved agricultural practices. It's important to remember that data quality and model selection (consider exploring advanced YOLO variants like YOLOv8) remain crucial for optimal performance. Additionally, continuous monitoring and refinement are essential to maintain effectiveness in real-world agricultural settings with varying conditions.

B. Convolutional Neural Networks for Automated Disease Detection in Agriculture

Convolutional neural networks (CNNs) are revolutionising automated disease detection in agriculture. Yolo-Papaya is a unique approach presented by Al-Sagheer et al. (2022) [8] that identifies and classifies diseases in papaya fruit.

CNNs are a type of deep learning model that focuses on image recognition and classification. They accomplish this through a series of specialist layers: Convolutional layers use filters like tiny grids to look for specific patterns in an image. CNNs may use numerous filters to extract a wide range of information, from simple forms to complex textures. Pooling layers: Following feature extraction, pooling layers reduce data dimensionality by merging information from fewer areas. This helps to regulate model complexity and prevent overfitting.

Fully connected layers: Finally, the processed data is passed into these layers, which look like classic neural networks. They connect neurons across layers, allowing the network to learn more advanced features and make classifications based on the retrieved data. The technique of feature extraction and hierarchical learning enables CNNs to recognise complicated patterns in images, making them perfect for disease diagnosis in agriculture. The Yolo-Papaya approach by Al-Sagheer et al. (2022) [8] exemplifies this. Precision agriculture researchers expanded on this notion by creating a new object detection system for quick and precise papaya fruit disease detection. This system makes use of YoloV7, a cutting-edge deep learning model noted for its speed and accuracy, and improves its performance with a Convolutional Block Attention Module (CBAM).

Prediction \ Ground Truth	Papaya (Healthy fruit)	Anthracnose	Phytophthora Blight	Mechanic Damage	Chocolate Spot	Sticky_Disease	Physiologicalspot	Black spot	Scar
Papaya (Healthy fruit)	727								
Anthracnose		105		2	2				
Phytophthora Blight			20						1
Mechanic Damage		2		93					20
Chocolate Spot		4			220				
Sticky_Disease						16			
Physiological Spot							120	10	
Black spot							15	374	
Scar			1	12					340
False Negative	9				42	5	16	83	83
Total	736	111	21	107	264	21	151	467	444

Fig. 1. Confusion Matrix [8]

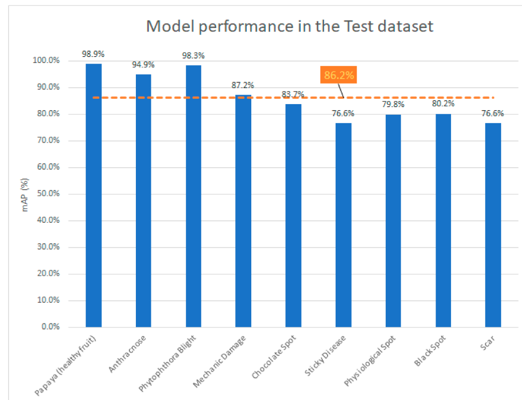

Fig. 2. Model Performace: A bar chart depicts class-specific model performance using mAP. A dashed line indicates the overall mean performance. [8]

The combined technique produced remarkable results, with an overall mean average precision (mAP) of 86.2%. The algorithm achieved 98.9% accuracy in classifying healthy fruits and those affected with Phytophthora blight, as illustrated in Fig. 2. This demonstrates the system's ability to distinguish between healthy and infected papayas, with a particular strength in detecting Phytophthora blight, a deadly fungal disease.

C. Studies

Deep learning methods are transforming automated disease detection in agriculture. Al-Sagheer et al. (2022) [9] present a new CNN architecture for detecting and classifying the severity of citrus crop diseases [9]. Their model uses a pre-trained VGG16 network, which is a well-known architecture

Class	Accuracy	Precision	Recall	F1 Score
Healthy	96%	100%	84%	91%
Low	99%	96%	100%	98%
Medium	97%	89%	100%	94%
High	98%	96%	96%	96%

Fig. 3. Confusion matrices for all levels of different disease in citrus crops. [9]

Class	Healthy	Low	Medium	High
Healthy	21	0	0	0
Low	0	25	0	1
Medium	3	0	25	0
High	1	0	0	24

Fig. 4. The accuracy, precision, recall, and F1 score of the model. [9]

for image recognition [9]. This pre-trained model serves as the system's foundation, allowing it to learn from enormous image collections. The authors then improve the VGG16 architecture to better categorise citrus disorders. The system can identify and diagnose diseases in citrus fruits based on labelled datasets with disease severity levels ranging from mild to severe.

The system learns to distinguish between healthy and diseased fruit by assessing visual features such as colour fluctuations, texture changes, and lesion patterns, and then applies these learned patterns to evaluate severity. This has several potential benefits, including early disease detection for rapid action, improved treatment techniques based on severity, and less reliance on physical inspection.

De Moraes et al. (2023) [8] present Yolo-Papaya, a deep learning model for papaya sickness classification that uses the YoloV7 detector with CBAM augmentation [8]. Both studies show that CNNs are effective for disease diagnosis, although their approaches differ. Paper [9] focuses on disease severity categorization for citrus fruits using a pre-trained VGG16 model, which could provide valuable insights into targeted treatment choices [9]. Future study might look at the models' applicability to different fruits, how they can be integrated into real-world agricultural applications, and how explainable AI techniques can help us understand how the models make their classifications.

III. Methodology

IV. Findings&Discussion of results

V. Conclusion

References

[1] "Climate change and agriculture in the United States: Effects and adaptation." [Online]. Available: https://www.researchgate.net/publication/269474231_Climate_change_and_agriculture_in_the_United_States_Effects_and_adaptation

[2] E. V. A. Dissanayake, *Research Onion: A Systematic Approach for Designing Research Methodology. Part One*, Feb. 2023.

[3] "What is Computer Vision? | IBM." [Online]. Available: https://www.ibm.com/topics/computer-vision

[4] M. Lippi, N. Bonucci, R. F. Carpio, M. Contarini, S. Speranza, and A. Gasparri, "A YOLO-Based Pest Detection System for Precision Agriculture," in *2021 29th Mediterranean Conference on Control and Automation (MED)*, Jun. 2021, pp. 342–347, iSSN: 2473-3504. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9480344

[5] R. Latha, G. Sreekanth, R. Rajadevi, S. Nivetha, K. Kumar, V. Akash, S. Bhuvanesh, and P. Anbarasu, "Fruits and Vegetables Recognition using YOLO," in *2022 International Conference on Computer Communication and Informatics (ICCCI)*, Jan. 2022, pp. 1–6, iSSN: 2329-7190. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9740820

[6] S. Alexandrova, Z. Tatlock, and M. Cakmak, "RoboFlow: A flow-based visual programming language for mobile manipulation tasks," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, May 2015, pp. 5537–5544, iSSN: 1050-4729. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7139973

[7] "Fruits and Vegetables Recognition using YOLO | IEEE Conference Publication | IEEE Xplore." [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9740820

[8] J. L. de Moraes, J. de Oliveira Neto, C. Badue, T. Oliveira-Santos, and A. F. de Souza, "Yolo-Papaya: A Papaya Fruit Disease Detector and Classifier Using CNNs and Convolutional Block Attention Modules," *Electronics*, vol. 12, no. 10, p. 2202, Jan. 2023, number: 10 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2079-9292/12/10/2202

[9] P. Dhiman, V. Kukreja, P. Manoharan, A. Kaur, M. M. Kamruzzaman, I. B. Dhaou, and C. Iwendi, "A Novel Deep Learning Model for Detection of Severity Level of the Disease in Citrus Fruits," *Electronics*, vol. 11, no. 3, p. 495, Jan. 2022, number: 3 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2079-9292/11/3/495