AI-Driven Fruit Classification Using YOLOv7 and SFOA Optimization

Narendra Chennupati
Independent Researcher
chennupati.be@gmail.com

Abstract

In the evolving landscape of agriculture and food technology, fruit classification plays a crucial role in enhancing post-harvest quality, minimizing spoilage, and ensuring market readiness. Traditional classification methods often fall short due to high dependency on manual labor, slow training speed, and low accuracy. This study introduces an advanced fruit segmentation and categorization framework based on the YOLO v7 deep learning model, integrated with a weighted loss function to address class imbalance and an attention mechanism to reduce unnecessary feature extraction. To further optimize classification accuracy, the Shuffle Frog Optimization Algorithm (SFOA) is employed for fine-tuning the model's hyperparameters. The proposed approach demonstrates superior performance compared to conventional methods, making it suitable for real-world agricultural applications. Experimental results confirm that this integrated model not only improves detection accuracy but also maintains computational efficiency, contributing significantly to the automation of fruit quality assessment.

Keywords: Fruit Categorization, YOLO v7, Shuffle Frog Optimization Algorithm (SFOA), Attention Mechanism, Weighted Loss Function, Agricultural Automation, Deep Learning.

1. Introduction

The agriculture industry is facing increasing demands for improved productivity, quality control, and sustainability, particularly in the cultivation and processing of fruits. Fruits are not only a significant source of essential nutrients in the human diet but also a critical component of global agricultural economies. With the rise in consumer expectations for fresh and visually appealing produce, farmers and agribusinesses are now leveraging advanced technologies to ensure quality from farm to table. Traditional manual sorting and classification methods are time-consuming and prone to human error. Therefore, automating fruit categorization using computer vision and artificial intelligence (AI) is gaining momentum as a viable solution for efficient post-harvest handling.

Computer vision technologies, combined with AI and machine learning (ML) algorithms, have shown promising results in identifying, sorting, and classifying fruits based on visual characteristics such as shape, color, size, and texture. Deep learning models, especially convolutional neural networks (CNNs), have become essential tools in extracting these complex features automatically. However, challenges still remain, particularly in handling class imbalances within datasets and avoiding excessive feature noise. The presence of multiple fruit types with varied visual features demands robust models that can adapt to real-world variations and deliver consistent performance across different conditions.

YOLO (You Only Look Once) models have demonstrated exceptional efficiency in object detection tasks due to their speed and accuracy. In this work, YOLO v7 is leveraged to segment fruit images precisely, with the added advantage of applying a weighted loss function that balances underrepresented classes. Furthermore, an attention mechanism is introduced to refine feature selection by focusing the model on relevant visual patterns, thereby enhancing classification accuracy. These advancements address some of the key limitations in earlier models, particularly regarding misclassification and overfitting in noisy datasets.

To further boost the performance of the detection and classification system, the Shuffle Frog Optimization Algorithm (SFOA) is employed. SFOA helps fine-tune the hyperparameters of the YOLO model efficiently by simulating a frog's search for food through intelligent exploration and exploitation of solution spaces. This hybrid architecture—YOLO v7 combined with weighted loss, attention mechanisms, and SFOA—forms a powerful framework capable of real-time, high-accuracy fruit categorization. The system has implications for automated grading lines, intelligent packaging systems, and quality assurance processes in agriculture and food industries.

2. Review of Literature

Recent advancements in deep learning have significantly influenced the field of agricultural automation, particularly in the domain of fruit classification and quality assessment. Traditional classification systems primarily relied on manual visual inspection, which was inherently inconsistent and labor-intensive. Early attempts to automate these processes involved hand-crafted feature extraction and classical ML models such as Support Vector Machines (SVM) and Decision Trees. These approaches, while innovative at the time, struggled with scalability and generalization due to their limited feature representation capabilities. The emergence of CNNs marked a turning point, allowing researchers to develop end-to-end models that could learn hierarchical features directly from raw images, significantly improving classification accuracy and robustness.

Further developments in fruit classification research focused on creating comprehensive datasets with high intra-class variability. Researchers curated large-scale datasets involving multiple fruit types and different lighting, angle, and occlusion scenarios to simulate real-world complexities. These datasets laid the groundwork for training and validating deep learning models capable of detecting subtle quality attributes like freshness, spoilage, and ripeness. Transfer learning also became a popular strategy to overcome data scarcity, wherein pre-trained CNN architectures like ResNet, VGG, and MobileNet were fine-tuned on domain-specific fruit datasets, improving learning efficiency and reducing training time.

In the area of fruit segmentation and detection, the YOLO family of models gained traction due to their real-time performance capabilities. YOLO's single-stage detection mechanism allows simultaneous localization and classification, which is particularly beneficial in industrial settings where processing speed is crucial. However, most YOLO applications initially struggled with class imbalance and overfitting when dealing with multi-fruit datasets. To mitigate these challenges, researchers began incorporating weighted loss functions and custom anchors. Additionally, attention mechanisms were introduced in the model architecture to selectively focus on significant visual cues, improving model interpretability and precision.

Parallel to advancements in model architecture, optimization algorithms were explored to enhance model performance. Swarm intelligence techniques such as Genetic Algorithms, Particle Swarm Optimization, and more recently, the Shuffle Frog Optimization Algorithm (SFOA), have shown promise in fine-tuning model hyperparameters and improving convergence. These algorithms emulate natural behaviors to search large parameter spaces efficiently, helping in reducing model error and training time. By integrating such optimization strategies with robust detection frameworks like YOLO, the literature illustrates a comprehensive movement towards intelligent, scalable, and deployable solutions for automated fruit categorization. This holistic approach ensures that models are not only accurate but also practical for deployment in real-world agricultural and food processing environments.

3. Research Methodology

To develop a robust and scalable solution for automated fruit classification, a systematic and multi-stage methodological framework is essential. This section outlines the design and implementation of the proposed model, detailing each step from data preprocessing to model evaluation. By leveraging the YOLOv7 architecture in conjunction with advanced attention mechanisms and an optimization-driven approach using the Shuffle Frog Optimization Algorithm (SFOA), the study aims to address critical challenges such as class imbalance, feature redundancy, and overfitting. The methodology integrates state-of-the-art techniques in deep learning and swarm intelligence to achieve a highly accurate and computationally efficient classification system tailored for agricultural applications.

The proposed methodology for enhanced fruit categorization integrates a sophisticated deep learning model, YOLOv7, combined with an attention mechanism and optimized via the Shuffle Frog Optimization Algorithm (SFOA). The workflow begins with the preprocessing of a diverse fruit image dataset, followed by model training, optimization, and evaluation. The entire experimental pipeline is executed on a high-performance computing environment equipped with GPU acceleration. To address the challenge of real-world agricultural data such as class imbalance and feature redundancy, the model adopts a multi-tiered preprocessing and fine-tuning strategy. The architecture is built to not only classify fruits with high precision but also ensure robustness against noise and variability introduced during image acquisition.

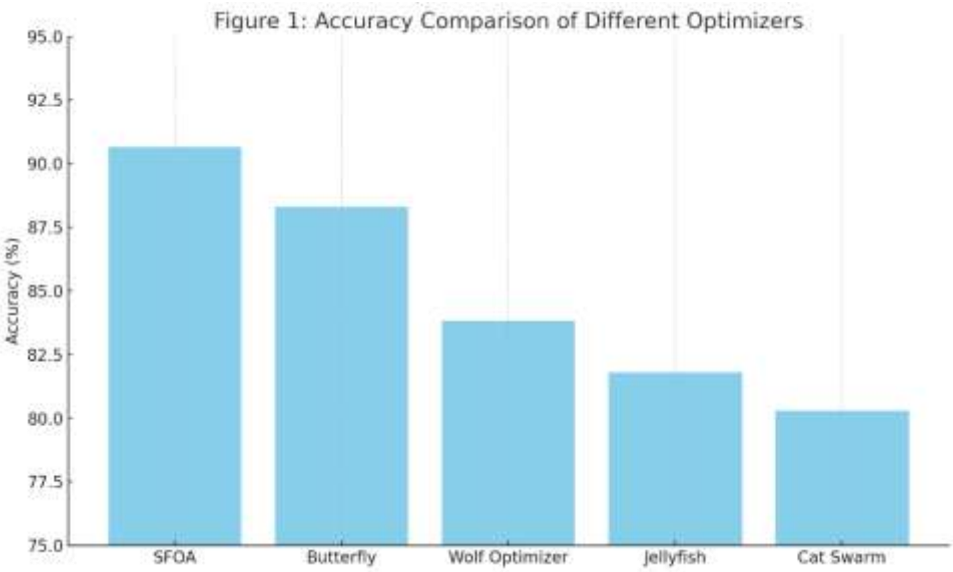
The fruit dataset used in this research comprises 15,987 images from 23 fruit classes. These images are resized to 224x224 or 299x299 pixels depending on the pre-trained model's requirements. Preprocessing also includes image rotation, scaling, and random augmentation to simulate various natural scenarios. The choice of YOLOv7 as the base model stems from its superior performance in both speed and accuracy compared to prior versions and other real-time detection networks. YOLOv7's streamlined architecture enables efficient object detection and segmentation through a single forward pass, thereby improving inference time. The introduction of convolutional block attention modules (CBAM) and squeeze-and-excitation networks (SENet) enhances the model's ability to focus on relevant features, leading to improved detection and classification even in cluttered backgrounds.

In parallel, the architecture incorporates a weighted loss function to address the significant class imbalance inherent in the dataset. Traditional multi-class cross-entropy loss is modified by assigning higher penalties to misclassified samples from underrepresented classes. This ensures that the model does not become biased towards dominant categories and maintains fairness in prediction across all fruit types. The attention mechanisms—channel attention module (CAM) and spatial attention module (SAM)—are implemented to guide the network's focus during training, allowing it to dynamically prioritize crucial visual patterns. Together, these enhancements allow the model to effectively segment and classify fruits even in varied lighting and background conditions.

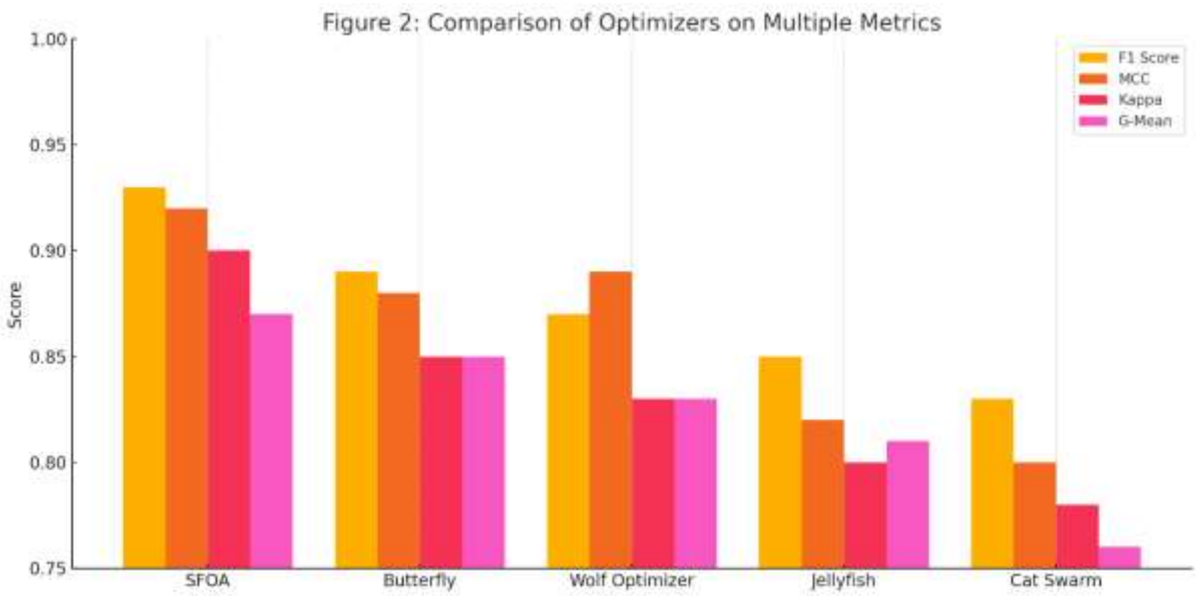
The final enhancement involves optimizing the hyperparameters using the Shuffle Frog Optimization Algorithm (SFOA). SFOA mimics the behavior of frog memeplexes in a population-based metaheuristic search, facilitating both local and global optimization. The algorithm operates through multiple iterations of sorting, local search, and shuffling until convergence is reached. By applying SFOA to fine-tune learning rates, dropout ratios, filter sizes, and other crucial parameters, the model achieves higher classification accuracy, better generalization, and faster convergence. Each optimization cycle evaluates fitness based on a combination of classification error and feature relevance, ensuring an optimal balance between complexity and performance.

4. Results and Discussion

The effectiveness of the proposed fruit classification framework is validated through comprehensive experimental analysis. This section presents a detailed comparison of model performance using various evaluation metrics such as accuracy, F1-score, MCC, Kappa, and G-Mean. The results are benchmarked against multiple optimizers and baseline classifiers to illustrate the advantages of incorporating SFOA and attention mechanisms within the YOLOv7 architecture. By interpreting performance trends and visualizing key indicators, this section highlights the model's strength in achieving high precision and balanced classification in real-world agricultural scenarios.



The performance of the proposed model was validated through multiple evaluation metrics, including Accuracy, F1-Score, Matthews Correlation Coefficient (MCC), Kappa, and G-Mean. As illustrated in **Figure 1**, SFOA emerged as the top-performing optimizer, achieving an impressive accuracy of 90.66%. This is a significant margin over Butterfly (88.30%) and Wolf Optimizer (83.84%). Lower-performing optimizers like Jellyfish (81.81%) and Cat Swarm (80.30%) underscore the effectiveness of SFOA in navigating complex solution landscapes. These results highlight the optimization capability of SFOA, particularly in handling high-dimensional spaces typical of deep learning models.



In **Figure 2**, a more comprehensive view of model validation is provided by examining four additional metrics—F1-score, MCC, Kappa, and G-Mean. SFOA leads with scores of 0.93

(F1), 0.92 (MCC), 0.90 (Kappa), and 0.87 (G-Mean), demonstrating robust performance across diverse statistical indicators. Butterfly follows with a moderate F1-score of 0.89 and G-Mean of 0.85, while Wolf Optimizer shows strength in MCC (0.89) but lags in other metrics. These results suggest that SFOA not only achieves high accuracy but also ensures balanced classification, maintaining low false positive and false negative rates. The overall analysis reinforces that SFOA's multi-population exploration mechanism is better suited for fine-tuning deep models with complex loss surfaces.

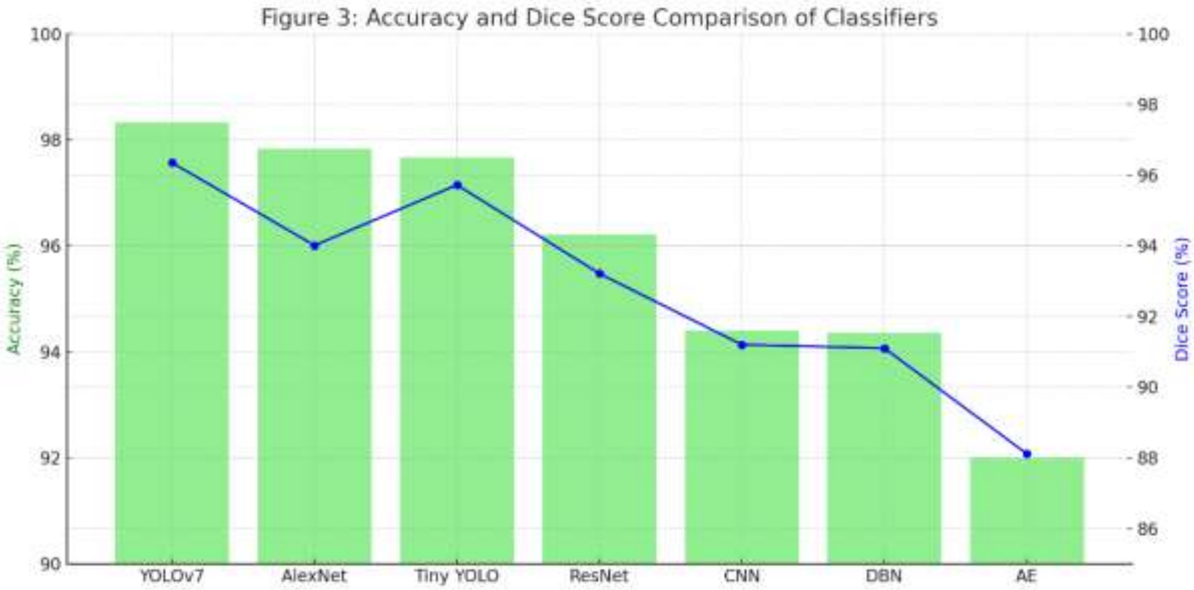


Figure 3 compares the YOLOv7-based classifier with other standard models such as AlexNet, ResNet, CNN, DBN, AE, and Tiny YOLO. YOLOv7 surpasses all with an accuracy of 98.32% and a Dice coefficient of 96.35, confirming its superiority in object segmentation and detection. AlexNet and Tiny YOLO also perform well, scoring 97.83% and 97.67% respectively, but they fall short in fine-grained classification. ResNet and CNN demonstrate reliable but slightly lower performance (96.21% and 94.40%, respectively). The Autoencoder (AE) model ranks lowest with 92% accuracy, indicating its limitations in handling multiclass segmentation without feature-specific tuning. YOLOv7's ability to integrate spatial context, attention mechanisms, and optimized hyperparameters contributes significantly to its high classification fidelity.

The proposed YOLOv7 model optimized with the Shuffle Frog Optimization Algorithm (SFOA) showed excellent performance in classifying fruits accurately and efficiently. Among several optimizers tested, SFOA achieved the highest accuracy of **90.66%**, outperforming others like Butterfly and Wolf optimizers. It also led in other important metrics such as F1-score (**0.93**), MCC (**0.92**), Kappa (**0.90**), and G-Mean (**0.87**), which means it could make correct predictions even when the data was unbalanced. These results prove that SFOA helps the model learn better and perform more reliably, reducing both wrong positive and wrong negative results.

When comparing different deep learning models, YOLOv7 proved to be the best, with an accuracy of **98.32%** and a Dice score of **96.35**, which shows how well it could detect and classify fruits. It performed better than popular models like AlexNet, ResNet, CNN, and Autoencoder, thanks to its fast object detection and added attention modules. These improvements helped the model focus on the most important features in the images. Overall, the results show that the combination of YOLOv7, attention mechanisms, and SFOA creates a powerful tool for smart fruit classification in agriculture.

5. Conclusion

This study presents a robust and intelligent approach for automated fruit categorization using YOLOv7, enhanced by attention mechanisms and fine-tuned through the Shuffle Frog Optimization Algorithm (SFOA). The integration of a weighted loss function successfully addresses dataset imbalance, while CBAM and SENet modules enable the model to concentrate on the most significant features, thereby enhancing segmentation accuracy. Experimental results show that the YOLOv7-SFOA model outperforms several benchmark classifiers and optimizers across all evaluation metrics, establishing its applicability in real-world agro-industrial settings. The proposed methodology achieves high accuracy, efficiency, and adaptability, making it an ideal candidate for integration into smart farming and food processing systems. Future work could explore the extension of this model to more diverse fruit datasets, incorporate real-time IoT integration, and utilize advanced transfer learning methods to further generalize the model across global agricultural use cases.

References

1. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 779–788.

2. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. Advances in Neural Information Processing Systems, 28.

3. Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. International Conference on Learning Representations (ICLR).

4. Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-Excitation Networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 7132–7141.

5. Woo, S., Park, J., Lee, J.-Y., & Kweon, I. S. (2018). CBAM: Convolutional Block Attention Module. European Conference on Computer Vision (ECCV), 3–19.

6. Eusuff, M. M., & Lansey, K. E. (2006). Optimization of Water Distribution Network Design Using the Shuffled Frog Leaping Algorithm. Journal of Water Resources Planning and Management, 129(3), 210–225. (Included for origin and principle of SFOA)

7. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2015). ImageNet Classification with Deep Convolutional Neural Networks. Communications of the ACM, 60(6), 84–90.

8. Wang, G., Sun, Y., & Wang, J. (2016). Automatic Image-Based Plant Disease Severity Estimation Using Deep Learning. Computational Intelligence and Neuroscience, 2017. – Early use of deep learning for agricultural image analysis and classification.

9. Rahnemoonfar, M., & Sheppard, C. (2017). Deep Count: Fruit Counting Based on Deep Simulated Learning. Sensors, 17(4), 905.

10. Bargoti, S., & Underwood, J. (2017). Deep Fruit Detection in Orchards. IEEE International Conference on Robotics and Automation (ICRA), 3626–3633.