ITM Web of Conferences **73**, 03016 (2025) https://doi.org/10.1051/itmconf/20257303016
*IWADI 2024*

# A Review of Optimization Strategy-Based Image Processing Techniques in Apple Picking

*Guoqin* Ma*

College of Humanities & Information Changchun University of Technology, Computer Science and Engineering, 130000, Changchun, China

**Abstract.** This paper reviews the application of optimization strategy-based image processing techniques in apple harvesting, focusing on the important role of various image processing methods in automated agriculture. These image processing methods include convolutional neural networks (CNN), support vector machines (SVM), artificial neural networks (ANN), and deep learning (DL) techniques. In apple target detection, localization, segmentation, ripeness, and quality detection. These techniques can effectively improve the accuracy and efficiency of automated picking systems, reduce the dependence on labor, and at the same time reduce the risk of fruit damage. However, there are still many challenges in practical applications, such as environmental factors like fruit occlusion, light variations, and complex backgrounds, which pose requirements on the adaptability and robustness of image processing techniques. To address these challenges, this paper explores various optimization strategies to improve the performance in complex environments. This paper looks forward to future research directions and proposes that the lightweight model should be further optimized, the real-time and adaptability of the algorithms should be improved, and low-cost and high-efficiency sensors and data processing techniques should be developed in order to promote the application and popularization of image processing technology in large-scale automated apple picking.

## 1 Introduction

As one of the wide fruits, harvesting apples globally requires a lot of labor every year. However, many challenges in apple harvesting include increasing production costs and limiting yields. Firstly, apple picking is highly dependent on labor [1]. There is much room for improvement in the efficiency of apple harvesting. Since there are differences in the skills of different pickers, the picking process may damage the apples' skin. To address these challenges in traditional harvesting, automation technology has been gradually introduced into the agricultural field. Image processing technology achieves accurate identification of apples [2]. This technology reduces the dependence on manpower and significantly improves picking efficiency. Image processing technology can recognize fruits in complex orchard environments. This technology also accurately detects apples through deep learning and

* Corresponding author: mgq@ldy.edu.rs

neural network algorithms. With the improvement of the optimization of algorithms, image processing technology in agriculture has become more and more promising [3]. First, this paper introduces several commonly used image processing techniques and their specific applications in the apple harvesting process. Second, this paper explores the challenges and optimization strategies of these techniques in practical applications. Finally, this paper also analyses the existing techniques' shortcomings and proposes the direction of future research and possible technological breakthroughs. This paper aims to provide reference and inspiration for the further development of automated apple picking systems in the future.

The other sections of this paper are shown. Section Two provides an overview of the history of image processing technology in agriculture, focusing on the background and needs of its application in automated apple picking. Section Three introduces the main image processing methods in detail, covering the core technologies of target detection, fruit localization, fruit ripeness, and quality detection, especially the detection methods based on deep learning. Section Four focuses on optimization strategies for image processing techniques, discussing how to improve the performance of apple-picking systems through algorithm optimization, data processing optimization, and environment adaptation optimization. Section Five analyses typical optimization application cases, which demonstrates the effectiveness of these techniques in actual picking, and discusses the challenges faced by the current techniques and the future development direction. Finally, Section Six provides an outlook on the future research direction of image processing techniques.

## 2 Critical Image Processing Techniques in Apple Harvesting

With the development of agricultural automation, image processing technology plays a crucial role in apple harvesting. A variety of image processing techniques are widely used in different aspects of apple harvesting, among which Convolutional Neural Networks (CNN), Support Vector Machines (SVM), Artificial Neural Networks (ANN), and Deep Learning (DL) are the most crucial types of techniques [4]. This paper details these techniques in apple harvesting and compares their advantages and disadvantages.

### 2.1 Convolutional Neural Network

CNN model can effectively extract features in images. CNN extracts local features of an image through multi-layer convolutional operations, which integrate these features through fully connected layers to achieve the classification of targets. Fan et al have shown that defect damage problems on the surface of apples [5]. Their approach can be effectively detected by convolutional neural networks with an accuracy as high as 93.74%.

### 2.2 Support Vector Machine

The SVM model can find the optimal separating surface to classify samples [6]. In apple harvesting, SVM is commonly used for fruit classification. In size and color classification of apples, SVM has achieved an accuracy of 98.17% [7]. This result is effective for orchard managers to select appropriate fruits. In addition, SVM also performs well in the automatic classification of apple size. This model can help the robot select the most suitable fruit. An SVM model performs well in classification tasks, but it may face several computational pressures [8]. Thus, researchers often combine principal component analysis (PCA) to reduce the data dimensionality. This combination approach may improve the computational efficiency.

### 2.3 Artificial Neural Networks

ANN is an algorithm that mimics the work of neurons [9]. In apple harvesting, An ANN model is widely used for defecting recognition, color analysis, and size determination of apples. In image defect detection, the accuracy of ANN is 91% [10]. Compared with traditional image processing algorithms, ANN has a stronger learning ability. In addition, ANN can automatically learn the color and size features of apples from the input image [11]. However, the training process of an ANN model usually requires numerous computational resources. Thus, researchers often combine migration learning techniques to improve training efficiency. Migration learning techniques can reduce the reliance on large-scale datasets.

### 2.4 Deep Learning

In apple harvesting, deep learning is widely used to solve occlusion detection [12]. These techniques include feature pyramid networks and region proposal networks in convolutional neural networks. These networks can accurately locate and segment occluded fruits through feature maps of different scales. In addition, deep learning plays an important role in Apple's 3D shape recognition. By combining multi-view images and 3D model techniques, deep learning algorithms can accurately reconstruct and identify the shape of apples. This is important for improving the efficiency of automatic apple picking and reducing fruit damage. Compared with traditional 2D image processing techniques, deep learning can better capture the spatial information of apples and help the picking robot to accurately locate the fruit gripping position, thus improving the success rate of picking [13]. Table 1 summarizes the application of different apple varieties in image processing techniques, listing the algorithms used for each apple object, their main features, and detection accuracies.

**Table 1.** Summarizes the application of different apple varieties.

| Object | Feature | Algorithm | Accuracy | Reference |
|---|---|---|---|---|
| Apple | Color, size, and shape | ANN | 95.00% | [12] |
| Apple | Texture and color | CNN and SVM | 93.00% | [13] |
| Apple | Color, size, and shape | Multi-spectral imaging | 92.00% | [14] |
| Red Fuji apple | Texture and color | CNN | 97.84% | [15] |
| Apple | Texture and color | SVM | 94.00% | [16] |
| Red Fuji apple | Texture and color | K-means | 90.00% | [17] |
| Red Fuji apple | Color, size, and shape | FuzzyNN | 94.00% | [18] |

Apple harvesting environment is complex and variable, different light conditions, weather conditions, and fruit density will affect the accuracy of image processing. Thus, how to improve the adaptability of image processing techniques and optimize algorithms to adapt to various complex environments has become a hot issue in current research. Table 2 shows an overview of the application of image processing algorithms in automated apple harvesting technology. This table summarizes the application of a variety of image processing algorithms to automated apple picking. It lists the name of each algorithm, the method of image capture, the image features extracted, and the apple features.

**Table 2.** An overview of the application of image processing algorithms.

| Algorithm | Image capturing approach | Image features | Apple features | Success rate | Picking cycle (sec) | Remarks |
|---|---|---|---|---|---|---|
| M. Henila, & P. Chithra [3] | monocular camera | Color segmentation, edge detection | Color (red, green), size | 85% | 10 | Better results in daytime conditions. |
| W. Jia et al. [9] | Multi-spectral imaging | Texture, color features | Multi-color fruit detection | 88% | 15 | Good fruit detection for occluded or shaded areas. |
| XD. Ma et al. [14] | RGB camera | Color threshold segmentation | Apple color and location detection | 96% | 9.5 | Stable performance under uneven lighting conditions. |
| N. M. Nasrabadi et al. [19] | Depth camera | Outline, edge features | Shape, depth information | 85% | 13 | Performs well in detecting fruit. |
| J. Rakun et al. [20] | Laser, RGB camera | 3D shape, surface reflection features | Shape, texture | 90% | 14 | Advanced hardware for high-precision. |
| L. Qiang et al. [21] | RGB, NIR camera | Shape, color, texture | Color and size detection | 88% | 12 | Performs well in combination with auxiliary lighting. |
| H. Song et al. [22] | Binocular camera | Edge detection | Overlapping fruits, depth information | 94% | 12 | Handles dynamic, swaying fruits effectively. |

| | | | | | | |
|---|---|---|---|---|---|---|
| Jia et al. [23] | RGB, Depth camera | Depth and color combined | Shape, ripeness | 92% | 14 | Optimizing apple detection. |

A review of CNN, SVM, ANN, and deep learning techniques shows that image processing techniques play an integral role in apple harvesting. Each algorithm has its strengths in different tasks; CNN excels in target detection and defect recognition, SVM has high accuracy in classification tasks, ANN has strong learning capabilities in dealing with complex image features, and deep learning can cope with complex tasks such as occlusion detection and 3D recognition [14]. Although these techniques have made some progress in the field of apple picking, they still face many challenges in practical applications. Future research should continue to explore how to further improve the efficiency and accuracy of automated apple picking by optimizing the algorithms and improving the robustness of the system.

# 3 Challenges in Apple Image Processing

Image processing techniques have been made the process to automated apple picking, but the complexity of the orchard environment makes the technique face many challenges. The main challenges include shading, changing light conditions, and the complexity of the outdoor environment. These challenges are summarized in the following, and relevant literature is cited to analyse and compare the effectiveness of different coping strategies. Image processing techniques provide strong support for fruit detection and picking during automated apple harvesting [21]. However, image processing faces a series of challenges due to the complexity of the orchard environment. In particular, how to ensure the robustness and accuracy of the system under the complexities of occlusion, light variations, and outdoor environments has become a pressing issue. This chapter will explore in detail the main challenges faced in apple image processing and present existing coping strategies and solutions.

## 3.1 Handling the Occlusion Problem

Occlusion is a major challenge in apple harvesting, and other fruits in orchard environments, making it difficult to accurately recognize occluded fruits with conventional image processing techniques. For example, the approach of Nasrabadi et al. [19] has high recognition accuracy in unobscured environments, the accuracy of detection significantly decreases when it encounters fruit occlusion. In addition, several combination algorithms show advantages in the occlusion problem. Mao et al combine CNN with SVM to improve the accuracy of apple recognition (93%), but this combination method increases computational complexity and is difficult to use in real-time scenarios [21]. Tournier et al performs well in occlusion (94%) [4]. However, their reliance on data makes them less adaptable. The main reason for the occlusion problem is that the natural environment of the orchard is highly dynamic. Therefore, several studies have consistently pointed out that the occlusion problem is difficult to solve due to its highly dynamic nature. However, these methods are still limited by computational complexity [4, 13, 19].

### 3.2 Changes in Lighting Conditions

Several studies point out that apple harvesting robots under strong light environments experience dramatic changes [14, 19, 21]. The effects of lighting conditions on images are global and difficult to eliminate with simple image enhancement techniques. In addition, complex lighting in orchards produces irregular brightness variations. The study suggests high dynamic range imaging techniques to address these problems. These techniques still leave much to be desired in terms of real-time and computational efficiency.

In contrast, other studies have proposed lighting normalization strategies, for example, Li & Lu suggested mitigating the effects of lighting variations through image preprocessing [13]. Although this approach is advantageous in reducing computational complexity, its effect is often less significant than HDR techniques. Overall, the HDR technique is superior in terms of accuracy, but the light normalization strategy requires less computational resources in practical applications. Different techniques have their advantages and disadvantages, and specific applications need to make a trade-off between real-time and accuracy. The multispectral imaging method showed some advantages in dealing with illumination changes, with an accuracy of 92%, and its performance was stable in the experimental environment [19]. However, the hardware cost of multispectral imaging is high and it is difficult to widely deploy in large-scale orchards

### 3.3 Complexity of Outdoor Environment

The complexity of outdoor environments increases the difficulty of distinguishing fruits from the background. In orchard environments, the color and textures of fruits and backgrounds are often highly similar, resulting in color segmentation-based algorithms performing poorly in complex backgrounds [24]. In addition, the small color and texture differences between leaves, branches, and fruits further increase the complexity of the detection system. To this end, Mao et al proposed the combination of CNN and SVM to improve the recognition accuracy in complex backgrounds by extracting the color and shape features of apple and successfully increased the accuracy to 93% [23]. However, this scheme still faces the problem of high computational demands in practical applications. The k-means algorithm achieves 90% accuracy by preliminary segmentation of fruit and background in complex background, but its low segmentation accuracy makes it difficult to be widely used in orchard environments [16]. In contrast, ANN shows certain advantages in processing color, size, and shape features, with an accuracy of 95%, but its high dependence on data and applicability is limited to some extent [14]. In conclusion, apple recognition in complex backgrounds is still a difficult problem in image processing. Deep learning algorithms perform well in this scenario, but their high demand for computational resources limits their wide application. Future studies can further optimize these algorithms to improve their performance in low-resource Settings.

## 4 Optimization Strategies for Image Processing Techniques

To address the image processing challenges, researchers have proposed combination algorithms, multimodal data, illumination optimization, and deep learning techniques. These strategies are summarized as follows.

### 4.1 Combining Different Algorithms to Improve Recognition Accuracy

A single algorithm may perform insufficiently in dealing with complex orchard environments. Mao et al successfully improved the accuracy of fruit classification by using CNNs to extract the color and shape features of apples [23]. This approach exploits the powerful feature of CNN and the efficient classification performance of SVM to solve the problem of the low recognition rate of traditional algorithms in complex backgrounds. However, the excellent performance of this hybrid strategy in experimental settings, and its effectiveness in practical applications remains controversial. The computational complexity of CNN is high, and the requirement of real-time processing is difficult to meet [17]. SVM needs to construct multiple classifiers when dealing with multi-category classification tasks. These tasks increase the complexity of the system [4]. Thus, these combination strategies can theoretically improve recognition accuracy. Their robustness and scalability in large-scale applications still need to be further verified. In contrast, although the single algorithm has low computational resource requirements, its performance is not as good as the combined strategy in complex environments.

### 4.2 Strategies for Applying Multimodal Data

The multimodal data effectively alleviates the limitations of traditional 2D image processing under occlusion conditions, and the study has shown that its detection accuracy has improved by 15-20%. However, multimodal data faces several problems [5]. The high cost of multimodal data acquisition equipment may not be economically viable for large-scale orchard applications. Multimodal data requires numerous computational resources, and real-time performance is still a problem to be solved. Although this strategy has achieved good results in laboratory environments, the limitations of computational resources may affect its diffusion in practical applications.

Jia et al. [9] also showed that the combination of RGB and depth image capture helped to optimize the fruit detection effect, with a success rate of 92% in occluded and dynamic environments. In addition, Rakun et al. [20] used a combination of laser and RGB cameras to capture 3D shape and surface features, which further improved the identification accuracy of apples (90% success rate). In addition, hybrid algorithms can also effectively solve the occlusion problem. Multimodal data strategies perform well in improving recognition accuracy, but their promotion in practical applications is limited by hardware cost and computational resources. Future research can reduce the application cost of multimodal data by developing low-cost sensors and efficient data processing techniques, to improve its practical operability.

### 4.3 Optimization Strategies under Light Variations

Illumination variation is a key factor affecting the effect of image processing. Jia et al proposed a multispectral imaging method that can maintain a relatively stable detection accuracy (88% success rate) under different lighting conditions and is suitable for the detection of multi-colored fruits in occluded and shaded areas [25]. However, the hardware cost of multispectral imaging is high, and the image acquisition speed is slow. In contrast, Ma XD et al, using an RGB camera and color threshold segmentation, achieved stable recognition results (96% success rate) under uneven lighting conditions, but showed low adaptability in occlusion and complex light change scenes [14]. In contrast, although the illumination normalization technique alleviates the problem of uneven illumination to some extent, it still has the risk of failure in the face of drastic illumination changes [23]. Thus, the illumination optimization strategy performs well under experimental conditions, but its real-time performance and computational resource problems in practical applications are still difficult problems in the future. Future research directions should focus on light processing algorithms that can adapt to dynamic light changes in real environments.

### 4.4 Optimization Strategies for Deep Learning Algorithms

Deep learning algorithms have been widely used in image processing. Han et al used a generative adversarial network (GAN) for data enhancement to simulate data under different lighting conditions, thereby improving the recognition accuracy of deep learning models in complex environments and increasing the success rate of the system to 94% [15]. Song et al used the contour and edge feature detection method of depth camera to achieve an 85% recognition success rate in the case of dense leaf occlusion, but the real-time performance of this method is low in the high-dynamic environment [18]. Although deep learning technology performs well in dealing with complex tasks, it is highly dependent on computing resources, and it is difficult to balance real-time performance and generalization ability, especially in large-scale applications. Future research can optimize the efficiency of deep learning algorithms to reduce the dependence on computing resources and improve the operability of practical applications.

## 5 Conclusion

With the rapid development of agricultural automation, the potential of image processing technology in apple picking has been widely recognized. By combining advanced artificial intelligence technology, image processing effectively solves the problems of low efficiency and high labour cost in traditional picking methods. However, there are still some limitations and shortcomings in previous studies. First, many algorithms have poor real-time performance in dealing with large-scale orchard environments. Second, although the application of multimodal data improves detection accuracy. Further, the robustness and generalization ability of existing techniques is still limited when facing problems. For large-scale data, the performance of deep learning models may fluctuate significantly.

Based on these limitations, future research can be conducted in the following directions: first, engineers should develop lightweight deep learning models to reduce the computational cost and validate their effectiveness in real-world automatic apple picking. This can help improve the operational efficiency of the models and also explore the performance differences of different models in diverse orchard environments. Second, developers should combine multimodal data to research low-cost sensors and efficient data processing techniques. Specifically, they can explore how to combine the sensing technology in agricultural science with the latest algorithms in computer vision and machine learning to improve the accuracy of data collection and processing efficiency. In addition, the exploration of interdisciplinary algorithms is an important direction for the future. Engineers need to develop algorithms that can be adapted to complex orchard environments to improve the robustness and adaptability of the algorithms. For example, they can combine light models in environmental science with image enhancement techniques in computer vision to cope with variable picking environments. Another direction worth exploring is the innovative hypothesis of multidisciplinary integration. For example, researchers should study agroecology in conjunction with machine learning to explore the impact of ecological factors on the performance of automated harvesting systems. Through the above interdisciplinary research directions and exploratory hypotheses, the application of image processing technology in automatic apple picking will not only make breakthroughs in technology but also show greater innovation and prospects in practical applications.

## References

1. S. Rotz, E. Duncan, M. Small, J. Botschner, R. Dara, I. Mosby, & E. D. Fraser, The politics of digital agricultural technologies: a preliminary review. Sociologia ruralis, **59(2)**, 203-229. (2019)
2. L. C. Ngugi, M. Abelwahab, & M. Abo-Zahhad, Recent advances in image processing techniques for automated leaf pest and disease recognition–A review. Information processing in agriculture, **8(1)**, 27-51. (2021)
3. M. Henila, & P. Chithra, Segmentation using fuzzy cluster‐based thresholding method for apple fruit sorting. IET Image Processing, **14(16)**, 4178-4187. (2020)
4. J. D. Tournier, R. Smith, D. Raffelt, R. Tabbara, T. Dhollander, M. Pietsch, & A. Connelly, MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. Neuroimage, **202**, 116137. (2019)
5. S. Fan, X. Liang, W. Huang, V. J. Zhang, Q. Pang, X. He, & C. Zhang, Real-time defects detection for apple sorting using NIR cameras with pruning-based YOLOV4 network. Computers and Electronics in Agriculture, **193**, 106715. (2022)
6. T. N. Syed, I. A. Lakhiar, & F. A. Chandio, Machine vision technology in agriculture: A review on the automatic seedling transplanters. International Journal of Multidisciplinary Research and Development, **6(12)**, 79-88. (2019)
7. A. Taner, M. T. Mengstu, K. Ç. Selvi, H. Duran, Ö. Kabaş, İ. Gür, & N. E. Gheorghiță, Multiclass apple varieties classification using machine learning with histogram of oriented gradient and color moments. Applied Sciences, **13(13)**, 7682. (2023)
8. C. Zhang, J. Valente, L. Kooistra, L. Guo, & W. Wang, Orchard management with small unmanned aerial vehicles: A survey of sensing and analysis approaches. Precision agriculture, **22(6)**, 2007-2052. (2021)
9. W. Jia, Y. Zhang, J. Lian, Y. Zheng, D. Zhao, & C. Li, Apple harvesting robot under information technology: A review. International Journal of Advanced Robotic Systems, **17(3)**, 1729881420925310. (2020)
10. X. Gao, S. Li, X. Su, Y. Li, L. Huang, W. Tang, & M. Dong, Application of Advanced Deep Learning Models for Efficient Apple Defect Detection and Quality Grading in Agricultural Production. Agriculture, **14(7)**, 1098. (2024)
11. C. Chen, B. Li, J. Liu, T. Bao, & N. Ren, Monocular positioning of sweet peppers: An instance segmentation approach for harvest robots. Biosystems Engineering, **196**, 15-28. (2020)
12. P. Pathmanaban, B. K. Gnanavel, & S. S. Anandan, Recent application of imaging techniques for fruit quality assessment. Trends in Food Science & Technology, **94**, 32-42. (2019)
13. B. Li, & W. Lu, Application of image processing technology in the digital media era in the design of integrated materials painting in installation art. Multimedia Tools and Applications, **83(18)**, 54211-54228. (2024)
14. XD. Ma, G. Liu, W. Zhou, et al., Apple recognition based on fuzzy neural network and quantum genetic algorithm. Trans Chin Soc Agric; **44(12)**: 227–232. (2013)
15. Z. Han, M. Jian, & G. G. Wang, ConvUNeXt: An efficient convolution neural network for medical image segmentation. Knowledge-based systems, **253**, 109512. (2022)

16. U. Shafi, R. Mumtaz, J. García-Nieto, S. A. Hassan, S. A. R. Zaidi, & N. Iqbal, Precision agriculture techniques and practices: From considerations to applications. Sensors, **19(17)**, 3796. (2019)
17. W. Jia, Y. Tian, R. Luo, Z. Zhang, J. Lian, & Y. Zheng, Detection and segmentation of overlapped fruits based on optimized mask R-CNN application in apple harvesting robot. Computers and Electronics in Agriculture, **172**, 105380. (2020)
18. H. Song, C., Zhang, J., Pan, X. Yin, & Y. Zhuang, Segmentation and reconstruction of overlappedapple images based on convex hull. Transactions of the Chinese Society of Agricultural Engineering, **29(3)**, 163-168. (2013)
19. N. M. Nasrabadi, Hyperspectral target detection: An overview of current and future challenges. IEEE Signal Processing Magazine, **31(1)**, 34-44. (2013)
20. J. Rakun, D. Stajnko, & D. Zazula, Detecting fruits in natural scenes by using spatial-frequency based texture analysis and multiview geometry. Computers and Electronics in Agriculture, **76(1)**, 80-88. (2011)
21. L. Qiang, C. Jianrong, L. Bin, D. Lie, & Z. Yajing, Identification of fruit and branch in natural scenes for citrus harvesting robot using machine vision and support vector machine. International Journal of Agricultural and Biological Engineering, **7(2)**, 115-121. (2014)
22. W. Jia, S. Mou, J. Wang, X. Liu, Y. Zheng, J. Lian, & D. Zhao, Fruit recognition based on pulse coupled neural network and genetic Elman algorithm application in apple harvesting robot. International Journal of Advanced Robotic Systems, **17(1)**, 1729881419897473. (2020)
23. W. Jia, D. Zhao, X. Liu, S. Tang, C. Ruan, & W. Ji, Apple recognition based on K-means and GA-RBF-LMS neural network applicated in harvesting robot. Transactions of the Chinese Society of Agricultural Engineering, **31(18)**, 175-183. (2015)
24. T. Georgiou, Y. Liu, W. Chen, & M. Lew, A survey of traditional and deep learning-based feature descriptors for high dimensional data in computer vision. International Journal of Multimedia Information Retrieval, **9**, 135-170. (2020)
25. W. Mao, B. Ji, J. Zhan, X. Zhang, & X. Hu, Apple location method for the apple harvesting robot. In 2009 2nd International Congress on Image and Signal Processing. IEEE. October (2009), pp. 1-5