Lightweight Architecture for Fruit Quality Estimation in the Infrared Domain

Patricia L. Suárez and Angel D. Sappa

Abstract This paper presents an automated apple classification system using near-infrared (NIR) imaging and novel lightweight convolutional neural networks (CNNs) to distinguish between apples in rotten, good, and bruised condition. The methodology relies on processing NIR images to capture subtle differences in the texture and composition of apple surfaces, which enables precise classification through the proposed network architecture. The algorithm optimizes identifying characteristic patterns associated with each category, allowing for improved accuracy and speed. Comparative results show that this approach outperforms state-of-the-art methods in efficiency, with potential applications in automated sorting systems and quality control processes.

Keywords Apple classification · Near-infrared imaging (NIR) · Image processing · NIR-based classification

1 Introduction

The automatic classification of agricultural products has become increasingly important in recent years, as it plays a critical role in ensuring product quality and reducing food waste. Apples, one of the most widely consumed fruits globally, are particularly susceptible to various types of damage during harvesting, transportation, and storage. Manual classification of apple quality is not only time-consuming but also prone to human error, especially when dealing with large quantities. Consequently,

http://www.espol.edu.ec.

P. L. Suárez (✉) · A. D. Sappa
ESPOL Polytechnic University, Guayaquil, Ecuador
e-mail: plsuarez@espol.edu.ec

A. D. Sappa
e-mail: asappa@espol.edu.ec; asappa@cvc.uab.es

A. D. Sappa
Computer Vision Center, Barcelona, Spain

S. Roy et al. (eds.), *Innovations in Computational Intelligence and Computer Vision*, Lecture Notes in Networks and Systems 1699,
https://doi.org/10.1007/978-3-032-09825-2_3

automated systems that can accurately classify apples into these categories: rotten, good, and bruised are of great interest to the agricultural and food industries.

Convolutional neural networks (CNNs) have emerged as one of the most effective techniques for image classification tasks, demonstrating superior performance in recognizing complex patterns in image data [4, 15]. CNNs have been successfully applied to a wide range of applications, including object detection, medical image processing, and facial recognition, just to mention a few. In the agricultural domain, CNNs have shown great promise in tasks such as disease detection, crop monitoring, and fruit classification [3, 5, 13, 23]. Traditional image processing methods, which rely on handcrafted features like color, texture, and shape [8], are often limited by their inability to generalize across different environmental conditions and image variations. CNNs, on the other hand, automatically learn hierarchical features from the data, enabling more robust and accurate classification [14, 18]. Also, some approaches use other spectra to efficiently tackle the fruit classification [6].

In addition to the challenges posed by the visual variability of apples, another significant issue is the availability of sufficient training data. To address this, data augmentation techniques such as random rotations, flips, and shifts were employed to expand the training dataset artificially. Data augmentation not only improves the model's ability to generalize but also helps prevent overfitting, which is a common issue in deep learning models when working with limited data [21, 25].

The current work aims to classify apples into three categories: rotten, good, and bruised using a CNN-based approach. Classifying damaged fruit poses several challenges, mainly due to the subtle differences that exist between good, bruised, and completely rotten apples. These differences are often difficult to capture using traditional image processing techniques. CNNs, with their ability to learn from large data sets to find patterns and their robustness to noise and variation, offer a compelling solution to this problem.

The remainder of this paper is organized as follows. Section 2 provides a detailed overview of works related to fruit classification using CNNs, highlighting the most relevant studies. Section 3 describes the methodology used in current work, including data preprocessing, model architecture, and training procedures. Section 4 presents the experimental results, and Sect. 5 discusses the findings and potential future work.

2 Related Work

The application of convolutional neural networks to agricultural image classification tasks has been significantly focused on in recent years. This section provides a review of the state of the art in fruit classification, with particular attention to the use of CNNs for apple classification and related tasks driven by the necessity for efficient and automated agricultural practices.

In [19], a CNN is proposed to classify good and rotten fruits is presented, which achieves an accuracy of 98.23%. This model uses multiple convolutional and pooling layers, effectively demonstrating the superiority of deep learning in automating fruit classification over traditional methods. Building upon this groundwork, a study presented in [20] introduces a web-based system for fruit identification, utilizing deep learning algorithms. This approach not only enhances user accessibility but also substantially improves classification accuracy, illustrating the potential of CNN integration in practical applications for farmers and consumers alike.

Recent research has emphasized the importance of transfer learning to improve the performance of classification algorithms. The study presented at [2] proposes a method that employs transfer learning using classical architectures, such as AlexNet, achieving a 99% accuracy rate in classifying fruits as good or rotten. Their findings highlight that the use of pre-trained models enables high classification performance, while also reducing computational complexity, making this approach feasible for real-world applications.

A comprehensive review has been presented in [22], examining various deep learning models employed in fruit detection and classification. The prevalence of traditional machine learning methods until recent years is noted, but a shift toward deep learning has markedly enhanced feature extraction capabilities. The review underscores the challenges presented by variations in fruit shape, color, and texture, which CNNs are increasingly addressing effectively.

In a recent study [10], the authors explore feature fusion techniques using CNN networks to improve the classification accuracy of fruit types. The study demonstrates that combining features from different layers of a CNN model results in improved recognition rates, thereby increasing overall performance on complex classification tasks. The effectiveness of deep learning methods is further supported by a systematic review presented in [12], which evaluates various techniques for fruit recognition using CNNs. Their results confirm that these networks consistently outperform traditional methods in accuracy and speed, confirming the advantages of deep learning in agricultural applications.

In addition to the works presented above, there are some contributions focused on specific types of fruits and their unique characteristics. For instance, in [17] the authors use CNNs to identify diseases in fruits by image analysis, demonstrating how timely disease detection can significantly help farmers implement necessary interventions. Other work for the detection of apple leaf disease is presented in [26], where the authors introduce a method for apple disease detection, which impacts fruit quality and value proposing a deep learning-based approach for automated disease detection, using a newly developed dataset of apple tree images from Kashmir, enhanced through expert annotation and augmentation techniques. Also, in [24], the authors present a review of various deep learning-based approaches utilizing specialized feature extraction methods for disease identification in tomato, potato, and apple crops. For fruit quality assessment [1] presents a deep learning approach with CNNs, introducing the DLFRUIT-GUI tool, which provides efficient predictions of fruit quality across varieties and storage conditions, enabling rapid and accessible user monitoring.

Similarly, in [11], a deep learning framework is introduced for automatically detecting and classifying fruits and vegetables, even in complex real-world scenarios. This approach supports fruit sellers by accurately identifying and differentiating similar-looking products, reducing manual sorting errors, and enhancing operational efficiency. Innovative architectures, such as EfficientNet, have been applied to improve fruit recognition capabilities. The study presented in [9] illustrates how these architectures enhance accuracy while maintaining computational efficiency, rendering them suitable for real-time applications.

In conclusion, this review of the state-of-the-art highlights the progress in fruit classification technologies powered by CNNs. These advancements contribute to enhancing agricultural productivity and ensure food quality and safety through timely interventions and improved decision-making processes. The main contribution of this paper is to propose a lightweight CNN architecture for apple quality classification, referred to as APC-CNN. This model has been trained from scratch and can be easily integrated into on-the-edge devices.

3 Proposed Approach

The proposed method is designed to integrate near-infrared imaging with a convolutional neural network for the classification of apples as rotten, bruised, and good. This section presents an outline of dataset definition, preprocessing, and CNN model design.

3.1 Dataset

The dataset is composed of samples acquired using two distinct cameras: the MER2-160-227U3C for visible spectrum images and the Basler ace acA1300-60gmNIR equipped with a BP660 bandpass filter (640–680 nm) centered at 660 nm for near-infrared imaging. These cameras offer resolutions of 1280×960 pixels and 1280×1024 pixels, respectively. In the current work, only the near-infrared samples were utilized, as this spectral band provides enhanced sensitivity to internal structures and water content, thus improving classification accuracy. NIR imaging allows for the detection of early signs of decomposition and bruising—features often invisible in the visible spectrum. The dataset is organized into 600 samples for training, 100 for validation, and 60 for testing. Figure 1 illustrates the data acquisition system employed for dataset collection, and Table 1 displays the sample distribution used for the experiments. Figure 2 shows some samples of each category used in the experiments for the classification process (as mentioned above in the current work just NIR images are considered).

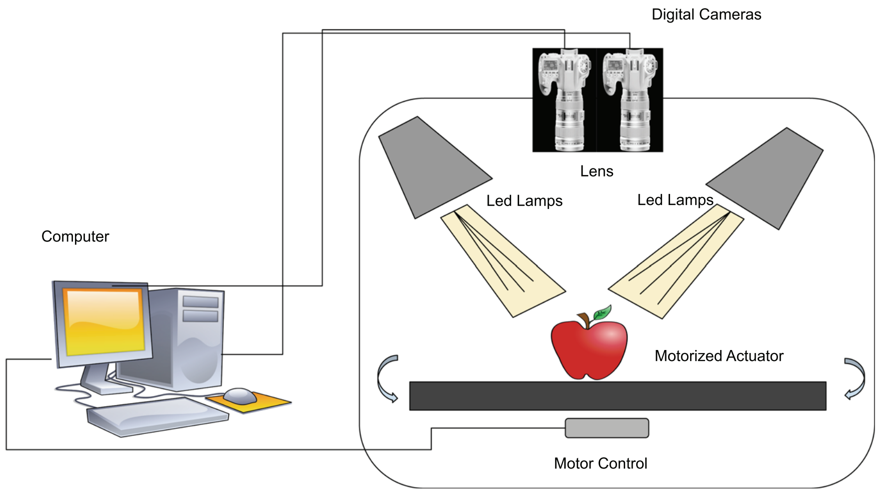


Fig. 1 Illustration of the acquisition system

Table 1 Dataset distribution used in apple classification

Dataset NIR images	Image distribution	Rotten instances	Good instances	Bruised instances
Training	600	170	200	230
Validation	100	30	30	40
Testing	60	22	18	20
Total	760	222	248	290

3.2 Data Preprocessing

To enhance the model's generalization capabilities, data augmentation techniques were implemented. Random rotations were applied to account for varying orientations of the apples, while horizontal and vertical flipping simulated symmetrical variations. Additionally, small changes in width and height were introduced to replicate slight shifts in apple positions. Finally, zooming and scaling were employed to mimic variations in camera distance and focus. These augmentations were applied dynamically during training to increase the diversity of the dataset and mitigate the risk of overfitting.

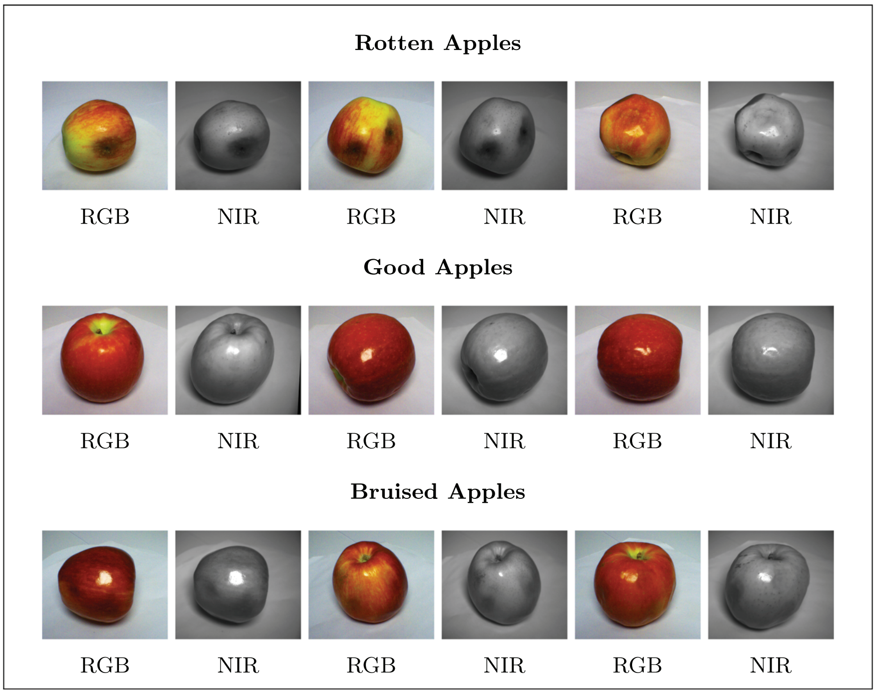


Fig. 2 Samples of the apple dataset

Table 2 Performance comparison on rotten, good, and bruised apple classification

Network	Rotten	Good	Bruised	Avg. accuracy	# of params (K)
VGG16	0.974	0.876	0.819	0.890	138,357
VGG19	0.994	0.896	0.839	0.910	143,667
ResNet50	0.986	0.860	0.931	0.925	25,636
ResNet101	0.996	0.880	0.951	0.942	44,549
APC-CNN	0.983	0.992	0.984	0.986	16,306

3.3 Loss Function

For this multi-class classification task, the categorical cross-entropy loss function was utilized. This function measures the divergence between the predicted probability distribution and the true labels. The categorical cross-entropy loss is defined as:

$$\mathcal{L}(y, \hat{y}) = -\sum_{i=1}^{N} y_i \log(\hat{y}_i) \quad (1)$$

where N is the number of classes, y_i is the true label (1 if the class is correct, 0 otherwise), and $\hat{y}_i$ is the predicted probability for class i.

3.4 Model Architecture

The CNN architecture has been designed to extract features from NIR images and classify the apples accordingly. The model begins with convolutional layers using a 3×3 filter to capture both low-level and high-level features, such as edges, textures, bruising, and internal decay. Max-pooling layers are incorporated with an initial pool size of 4×4 after the convolutional layers and followed by a 2×2 pool size in subsequent layers, with strides carefully selected to progressively reduce the dimensionality of the feature maps while retaining essential information and at the same time decreasing the model's computational complexity. This also helps to mitigate overfitting by removing extraneous details from the input. Multiple convolutional layers are stacked, with each successive layer using additional filters to capture increasingly complex features from the input data. The LeakyReLU activation function was applied after each convolutional layer to introduce non-linearity and prevent the dying ReLU problem, which can arise when neurons stop learning due to large negative inputs [16]. This function allows a small, positive gradient for negative inputs, enhancing the learning process, particularly when distinguishing subtle differences like bruising.

Following the feature extraction process, the output of the convolutional layers is flattened and passed through fully connected (dense) layers. The first dense layer comprises 1024 units and employs the ReLU activation function to integrate the learned features. The final dense layer consists of 3 units, corresponding to the three classes, with a softmax activation function applied to produce the probability distribution over the classes. Figure 3 illustrates the proposed architecture.

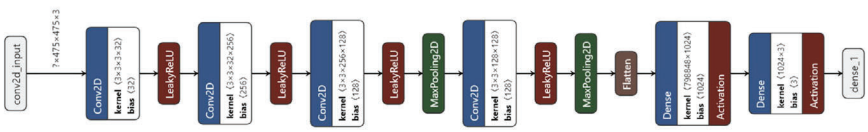


Fig. 3 Detailed classification model architecture

The model was trained with the Nadam optimizer, a variant of the Adam optimizer that incorporates Nesterov momentum to enhance convergence speed and accuracy [7]. The learning rate was initialized at 0.000032, and hyperparameters such as beta values were meticulously tuned to optimize model performance.

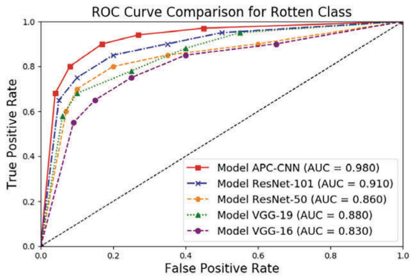
4 Experimental Results

This section presents the experimental results obtained with the proposed architecture. The model's classification performance is evaluated and compared with state-of-the-art approaches.

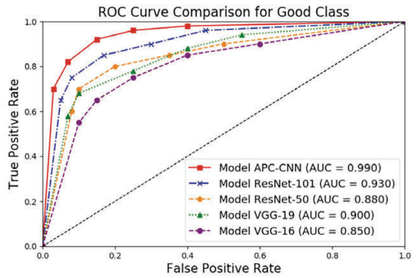
The proposed APC-CNN model was compared against other state-of-the-art models, including VGG16, VGG19, ResNet50, and ResNet101. The respective pre-trained networks were used in each experiment, and a fine-tuning process was applied to adapt them to our dataset. For VGG16, the first 14 layers were retained, while the final two layers were retrained to better capture features specific to our Apple dataset. For ResNet50 and ResNet101, the entire set of parameters was fine-tuned to fully utilize the capacity of these deeper networks, improving classification accuracy. Comparative results are shown in Table 2, indicating that the APC-CNN approach outperforms other state-of-the-art methods. APC-CNN achieves higher accuracy in classifying the three apple categories: rotten, good, and bruised.

A notable aspect of our approach is the significantly reduced number of parameters, compared to other methods. APC-CNN has the smallest parameter count among all the evaluated techniques, which reduces the computational cost and accelerates convergence during training. This efficiency renders APC-CNN suitable for real-time applications or scenarios with limited computational resources while still maintaining high classification accuracy.

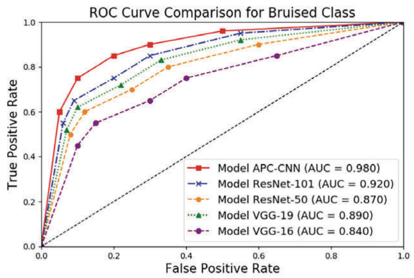
Another analysis conducted to evaluate the model's performance and accuracy involves generating the Receiver Operating Characteristic (ROC) curve, which is a graphical representation illustrating classifier performance by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings. Figure 4 provides a comparative analysis of the ROC curves for APC-CNN against VGG16, VGG19, ResNet50, and ResNet101. As shown, APC-CNN consistently produces the best ROC curve across all classes, showcasing its superior capability to distinguish between categories. In conclusion, the proposed APC-CNN model outperforms other state-of-the-art approaches. This makes it an efficient and effective model for apple classification tasks, offering both high precision and speed for practical applications.



(a) Rotten class



(b) Good class



(c) Bruised class

Fig. 4 ROC curves for comparison of rotten, good, and bruised classes with other state-of-the-art approaches

5 Conclusions

This paper presents a simple yet effective neural network model designed to classify apple quality into three categories: rotten, good, and bruised. The strength of the proposed approach lies in its simplicity and robustness, as it does not rely on additional preprocessing steps or external features to assess apple quality. Instead, the model processes the input images directly, using NIR images, which provide enhanced detection capabilities that further improve classification accuracy. The robustness of the approach is demonstrated through experimental results, which highlight its ability to accurately classify apples under various quality conditions without requiring extensive fine-tuning or additional parameters. These results showcase the model's capacity for generalization, maintaining high accuracy across diverse test scenarios.

Acknowledgements This work has been supported by the ESPOL project CIDIS-005-2024. The authors acknowledge the support of the Generalitat de Catalunya CERCA Program to CVC's general activities, and the Departament de Recerca i Universitats from Generalitat de Catalunya with reference 2021SGR01499.

References

1. M.M. Ali, N. Hashim, Development of deep learning based user-friendly interface for fruit quality detection. J. Food Eng. **380**, 112165 (2024)
2. U. Amin, M.I. Shahzad, A. Shahzad, M. Shahzad, U. Khan, Z. Mahmood, Automatic fruits freshness classification using CNN and transfer learning. Appl. Sci. **13**(14), 8087 (2023)
3. L. Chuquimarca, B. Vintimilla, S. Velastin, Banana ripeness level classification using a simple CNN model trained with real and synthetic datasets, in *Proceedings of the 18th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP)*, vol. 5 (INSTICC, SciTePress, 2023), pp 536–543. https://doi.org/10.5220/0011654600003417
4. L. Chuquimarca, B. Vintimilla, S. Velastin, Classifying healthy and defective fruits with a multi-input architecture and CNN models, in *14th International Conference on Pattern Recognition Systems (ICPRS)* (2024), pp. 1–7. https://doi.org/10.1109/ICPRS62101.2024.10677833
5. L.E. Chuquimarca, B.X. Vintimilla, S.A. Velastin, A review of external quality inspection for fruit grading using CNN models. Artif. Intell. Agric. **14**, 1–20 (2024). https://www.sciencedirect.com/science/article/pii/S2589721724000369
6. O. Coello, M. Coronel, D. Carpio, B. Vintimilla, L. Chuquimarca, Enhancing apple's defect classification: insights from visible spectrum and narrow spectral band imaging, in *14th International Conference on Pattern Recognition Systems (ICPRS)* (2024), pp. 1–6. https://doi.org/10.1109/ICPRS62101.2024.10677803
7. T. Dozat, Incorporating Nesterov momentum into Adam (2015). https://api.semanticscholar.org/CorpusID:620137
8. S.R. Dubey, A.S. Jalal, Apple disease classification using color, texture and shape features of images. Signal Image Video Process. **10**(5), 819–826 (2016)
9. L.T. Duong, P.T. Nguyen, C. Di Sipio, D. Di Ruscio, Automated fruit recognition using EfficientNet and MixNet. Comput. Electron. Agric. **171**, 105326 (2020)
10. H.S. Gill, G. Murugesan, A. Mehbodniya, G.S. Sajja, G. Gupta, A. Bhatt, Fruit type classification using deep learning and feature fusion. Comput. Electron. Agric. **211**, 107990 (2023)
11. D. Hussain, I. Hussain, M. Ismail, A. Alabrah, S.S. Ullah, H.M. Alaghbari, A simple and efficient deep learning-based framework for automatic fruit recognition. Comput. Intell. Neurosci. **1**, 6538117 (2022)
12. D. Indira, J. Goddu, B. Indraja, V.M.L. Challa, B. Manasa, A review on fruit recognition and feature evaluation using CNN. Mater. Today Proc. **80**, 3438–3443 (2023)
13. A. Kamilaris, F.X. Prenafeta-Boldú, Deep learning in agriculture: a survey. Comput. Electron. Agric. **147**, 70–90 (2018)
14. A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks. Adv. Neural Inf. Process. Syst. **25**, 1097–1105 (2012)
15. Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition. Proc. IEEE **86**(11), 2278–2324 (1998)
16. A.L. Maas, A.Y. Hannun, A.Y. Ng, Rectifier nonlinearities improve neural network acoustic models. Proc. ICML **30**, 3 (2013)
17. I.M. Nasir, A. Bibi, J.H. Shah, M.A. Khan, M. Sharif, K. Iqbal, Y. Nam, S. Kadry, Deep learning-based classification of fruit diseases: an application for precision agriculture. Comput. Mater. Contin **66**(2), 1949–1962 (2021)
18. R. Pacheco, P. González, L. Chuquimarca, B. Vintimilla, S. Velastin, Fruit defect detection using CNN models with real and virtual data, in *Proceedings of the 18th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP)*, vol. 4 (INSTICC, SciTePress, 2023), pp. 272–279. https://doi.org/10.5220/0011679800003417
19. R. Pathak, H. Makwana, Classification of fruits using convolutional neural network and transfer learning models. J. Manag. Inf. Decis. Sci. **24**, 1–12 (2021)

20. M.M. Rahman, M.A. Basar, T.S. Shinti, M.S.I. Khan, H.M.H. Babu, K.M.M. Uddin, A deep CNN approach to detect and classify local fruits through a web interface. Smart Agric. Technol. **5**, 100321 (2023)
21. P.L. Suárez, H.O. Velesaca, D. Carpio, A.D. Sappa, Corn kernel classification from few training samples. Artif. Intell. Agric. **9**, 89–99 (2023)
22. C.C. Ukwuoma, Q. Zhiguang, M.B. Bin Heyat, L. Ali, Z. Almaspoor, H.N. Monday, Recent advancements in fruit detection and classification using deep learning techniques. Math. Probl. Eng. **1**(1), 9210947 (2022)
23. H.O. Velesaca, P.L. Suárez, R. Mira, A.D. Sappa, Computer vision based food grain classification: a comprehensive survey. Comput. Electron. Agric. **187**, 106287 (2021)
24. J.A. Wani, S. Sharma, M. Muzamil, S. Ahmed, S. Sharma, S. Singh, Machine learning and deep learning based computational techniques in automatic agricultural diseases detection: methodologies, applications, and challenges. Arch. Comput. Methods Eng. **29**(1), 641–677 (2022)
25. L. Wu, J. Ma, Y. Zhao, H. Liu, Apple detection in complex scene using the improved YOLOv4 model. Agronomy **11**, 476 (2021)
26. A.A. Yatoo, A. Sharma, An indigenous dataset for the detection and classification of apple leaf diseases. Data Brief **53**, 110165 (2024)