www.editada.org

International Journal of Combinatorial Optimization Problems and Informatics, 16(3), May-Aug 2025, 106-117. ISSN: 2007-1558. https://doi.org/10.61467/2007.1558.2025.v16i3.1141

Application of Deep Learning for Automated Peach Classification: A Study Based on ResNet Architectures

Flossi Puma-Ttito[1], *Carlos Guerrero-Mendez* *[,1], *Daniela Lopez-Betancur* *[,1,2], *Tonatiuh Saucedo-Anaya*[1], *Rafael Castaneda-Diaz*[1,3], *Luis Martinez-Ytuza*[1]

[1] Universidad Autónoma de Zacatecas, Unidad Académica de Ciencia y Tecnología de la Luz y la Materia (LUMAT), Zacatecas, México.

[2] Universidad Autónoma de Zacatecas, Unidad Académica de Ingeniería, Zacatecas, México.

[3] Instituto Politécnico Nacional, Unidad Profesional Interdisciplinaria de Ingeniería Campus Zacatecas (UPIIZ), Zacatecas, México.

guerrero_mendez@uaz.edu.mx, danielalopez106@uaz.edu.mx

Abstract. This study evaluates the performance of various ResNet architectures for classifying peaches as "healthy" or "damaged". A dataset of 3 370 images was used, with data-augmentation techniques applied to enrich the training set. Transfer learning was performed using pre-trained ResNet models, with stochastic gradient descent (SGD) adopted as the optimisation algorithm. Performance was assessed using accuracy, precision, recall and F1 score. ResNet-50 emerged as the most effective architecture, achieving a mean accuracy of 95.96 % and outperforming other models, including ResNet-18, ResNet-34, ResNet-101 and ResNet-152. The results demonstrate the potential of deep-learning techniques to improve peach-sorting processes, thereby reducing post-harvest losses and enhancing quality control in the agricultural sector.

Keywords: Image classification, ResNet architecture, Deep Learning, Peach classification, Convolutional Neural Networks.

Article Info
Received May 23, 2025
Accepted Jul 1, 2025

1 Introduction

The peach (*Prunus persica L.*), a member of the Rosacea family, is enjoyed worldwide in various forms, from fresh and sliced to incorporated into preserves, syrups, and desserts. It is a good source of vitamins A, B1, B2, and C, along with minerals like phosphorus and calcium (Africano P. et al., 2015), the peach is nonetheless highly perishable due to its high-water content. Their short shelf life, transport and post-harvest storage can cause losses of between 15 and 25% of total production (Gonzales et al., 2022). Efficient peach sorting is paramount for delivering a high-quality product and ensuring consumer satisfaction. Existing manual sorting methods are often slow, labor-intensive, and prone to inaccuracies. Several factors contribute to peach quality, including size, color, and flavor, while the presence of physical damage, pests, diseases, or foreign matter designates a peach as damaged. Although standards such as the Mexican Official Standard NMX-FF-060-SCFI-2009 (Diario Oficial de la Federacion, 2009) specify quality requirements, traditional markets frequently lack objective grading, relying instead on the experience of individual sellers. Advanced sorting technologies offer peach marketers a promising avenue for modernizing and improving their operations.

A human brain exhibits sophisticated information processing capabilities, enabling problem-solving, decision-making, and evaluation of information derived from both external and internal sources (Corvalán, 2018; Flores et al., 2022). Artificial Intelligence (AI) seeks to emulate these cognitive functions by developing computer systems capable of learning from data through training (Naranjo-Torres et al., 2020). Applications of AI are diverse and span various fields, including education, science, and technology.

AI is now widely used to solve some human problems more efficiently. A subfield of AI, known as Deep Learning (DL), utilizes computational models with multiple processing layers to learn complex data representations at various levels of abstraction. These representations enable applications such as speech recognition, visual object recognition, and object detection. Convolutional Neural Networks (CNN) are particularly prominent within DL. CNNs are algorithms designed to mimic aspects of human learning through interconnected computational blocks and layers of artificial neurons that can approximate complex functions. These models excel at image analysis tasks, including classification and detection. The CNNs have made significant advances in processing images, videos, voice, and audio for classification or identification (LeCun et al., 2015). In agriculture, CNNs have been used for classification and detection of fruits (Naranjo-Torres et al., 2020), diseases (Maeda-Gutiérrez et al., 2020), sizes (Medina et al., 2022), colors (Méndez Almansa & Silva Salamanca, 2022), etc. The CNNs offer a promising solution due to their ability to automatically learn relevant features from image data. Using AI and DL algorithms, state-of-the-art sorting machines have been developed that analyze images to detect abnormal or defective products with high precision, without the need for human intervention (Lopez-Betancur, Saucedo-Anaya, et al., 2024; Navarro-Solís et al., 2024). These machines can distinguish between different samples of fruits based on characteristics such as size, color, weight and other product properties. The development of these technologies is of great interest to agricultural producers, as it allows them to classify their products more efficiently.

Peach classification is a crucial task, and researchers have explored various methods and techniques to address it. A recent study conducted at the Universidad Autónoma del Estado de México employed CNNs to classify peaches (Akbar et al., 2022). A dataset of 960 images was captured using a Nikon D3500 camera. This dataset included 360 images of ripe and unripe peaches and 600 images of healthy and damaged peaches. 80% of the dataset was used for training, while the remaining 20% was reserved for validation. The researchers proposed a straightforward CNN architecture that yielded a classification accuracy of 95.31% for distinguishing between ripe and unripe peaches and 92.18% for classifying healthy and damaged peaches.

Yao (Yao et al., 2022) applied DL models, specifically Mask R-CNN and Mask Scoring R-CNN, to segment and recognize various peach diseases, including brown rot, anthracnose, scab, bacterial shot hole, gummosis, powdery mildew, and leaf curl. The dataset comprised 94 images for brown rot, 157 for anthracnose, 654 for scab, 427 for bacterial shot hole, 91 for gummosis, 50 for powdery mildew, and 87 for leaf curl. By utilizing ResNet50 and ResNet101 as backbone networks for Mask R-CNN, the segmentation accuracy, as measured by segm_mAP_50, improved from 0.236 to 0.254 and from 0.452 to 0.463, respectively.

Akbar and other researchers investigated bacteriosis (Akbar et al., 2022), a significant disease affecting peach crops worldwide. In this paper, that proposed a lightweight CNN model, WLNet, based on the VGG-19 architecture to detect and classify peach leaf images as either bacteriosis-infected or healthy. The dataset used for training and testing consisted of 10,000 images, with 4,500 bacteriosis-infected and 5,500 healthy images. WLNet was compared to four other CNN models: LeNet, AlexNet, VGG-16, and a simple VGG-19 model. The proposed WLNet model achieved an accuracy of 99%, outperforming the other models. These results demonstrate the effectiveness of WLNet in accurately detecting bacteriosis in peach leaf images.

Based on previous research, this study is purpose to evaluate the performance of several advanced Torchvision architectures (an open-source computer vision package for the Torch machine learning library) for peach quality classification. By comparing five Torchvision models available in PyTorch (ResNet-18, ResNet-34, ResNet-50, ResNet-101 and ResNet-152), the goal is to identify the most effective model for potential implementation in a peach sorting machine. The performance of these models is assessed using cross validation and standard metrics of DL.

2 Methods, Techniques and Instruments

This section describes the use of five CNN models for peach image classification. The equipment, methods, and performance metrics used in this research are also detailed.

2.1 CNN models

CNNs are a class of artificial neural network architectures composed of blocks that work together to process images. The main function of their layers is to identify relevant features of an image (Taye, 2023). These architectures consist of stacked convolutional layers, and as technology advances, the architectures of CNNs also evolve, allowing them to tackle increasingly complex image-related tasks.

A CNN ResNet, also known as Residual Network, marked a significant breakthrough in the fields of DL and computer vision when it was introduced by He et al (He et al., 2016). The core idea of ResNet is to address a common problem in deep networks: the degradation that occurs during training. As a network becomes deeper, one would expect it to have a greater capacity to represent information, but in practice, the training error tends to increase. This is due to a problem known as the vanishing or exploding gradient, which makes deep networks difficult to train. ResNet solves this problem by introducing skip connections that allow signals to flow directly from early to later layers of the network, helping to maintain training stability, as shown in Figure 1.

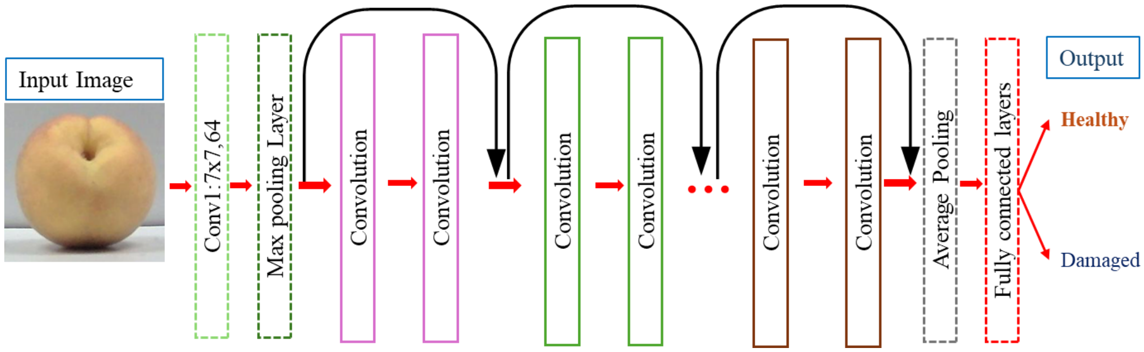

Fig. 1. A schematic view of ResNet architecture for classifying peach images.

The schematic diagram of the ResNet architecture illustrates that it is composed of an input layer, multiple stacked residual blocks (Res Block), and an output layer. Each residual block permits the direct transfer of features through skip connections, which assists in preventing the vanishing gradient problem in deep networks. This visual representation demonstrates how ResNet improves learning capacity and training efficiency by incorporating residual learning units. When the residual is equal to zero, the skip connections can transmit the input without altering the performance of the network. A distinctive characteristic of ResNet is its layered design.

In practice, ResNet employs Batch Normalization and the ReLU activation function to enhance training stability and convergence rate. Batch Normalization standardizes the data before each layer, mitigating internal covariate shift and enabling higher learning rates. The ReLU function adds nonlinearity, boosting the expressive power of the network.

Given these foundational elements, ResNet has been developed in various configurations to address different computational needs and performance requirements. Below, we outline the key characteristics of the most prominent ResNet variants:

• **ResNet-18**
The ResNet-18 model, with its 18 layers, offers a balanced trade-off between training time and computational efficiency, making it highly effective for image classification tasks.

• **ResNet-34**
With 34 layers, ResNet-34 is a versatile architecture suitable for various computer vision tasks. Its ability to capture complex features and identify objects in images positions it as a superior choice compared to ResNet-18.

• **ResNet-50**
Featuring 50 layers, ResNet-50 excels in learning highly complex image representations. It is widely used for large-scale image classification, object detection, and image segmentation, thanks to its robust performance and extensive support in DL libraries.

• **ResNet-101**
Composed of 101 layers, ResNet-101 is deeper than its predecessors. This depth allows it to capture more intricate patterns and features, making it particularly effective for challenging image classification tasks. However, its increased depth also demands higher computational resources.

• **ResNet-152**
With 152 layers, ResNet-152 is one of the deepest neural networks available. Its architecture enables it to learn extremely complex and abstract features, making it ideal for tasks requiring high precision. ResNet-152 has demonstrated exceptional performance in image classification, achieving very high accuracy on datasets like ImageNet. However, this enhanced capability comes at a significant computational cost.

ResNet is a prominent CNN architecture for image classification, utilizing deep networks and residual blocks to effectively extract high-level features and address training challenges in deep models. Known for its high accuracy and efficiency, ResNet stands out among CNN-based models, as reported by Torchvision documentation. In this study, various ResNet models (ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152) were evaluated using identical hyperparameters. The evaluation focused on performance metrics and processing time. The objective was to identify the most suitable architecture that achieves a balance among accuracy, complexity, and processing time for peach image classification.

2.2 Transfer Learning

Training a CNN model from scratch requires many labeled images. To reduce computational costs and training time, Transfer Learning (TL) is a highly utilized and recommended approach. This technique trains only a part of the pre-trained model to perform the task of classifying images (Alebiosu & Muhammad, 2019).

In this research, ResNet architecture with pre-trained weights on ImageNet will be used and retrained for peach classification. The network layers will be frozen, except for the final layer, which will be modified to have the two necessary outputs and trained using the peach database. Table 1 presents the total parameters of the CNN models and the parameters that will be trained after freezing the network. In addition, the disk space that each model occupies is detailed.

Table 1. Size on disk and total of trainable parameters in ResNet models.

CNN model	CNN with TL (Trainable parameters)	CNN pre-trained with ImageNet (Total parameters)	Size on disk (MB)
ResNet-18	1026	11689512	44.7
ResNet-34	1026	21797672	83.3
ResNet-50	4098	25557032	97.8
ResNet-101	4098	44549160	170.5
ResNet-152	4098	60192808	230.4

2.3 Data Acquisition

The Mexican Official Standard NMX-FF-060-SCFI-2009 establishes the requirements for peaches (Prunus persica L.) to meet for fresh marketing in Mexico, ensuring they are healthy, intact, free of pesticides, and possess the organoleptic characteristics of their variety. This standard provides clear guidelines on what constitutes a healthy peach and what indicates damage, which is crucial for the classification task of distinguishing between healthy and damaged peaches.

To create an image database, 316 peaches were collected from a local grower in Zacatecas, Mexico. Images were captured in a controlled environment with white light, using an Atvio A489 camera positioned at a fixed distance from each peach on a stationary white background, resulting in photographs of 420x400 pixels.

The final dataset comprised 3370 images, categorized into healthy peaches (2470 images) and damaged peaches (900 images). The dataset was split into 80% for training and 20% for validation, maintaining the class proportions. Figure 2 provides examples of each class.

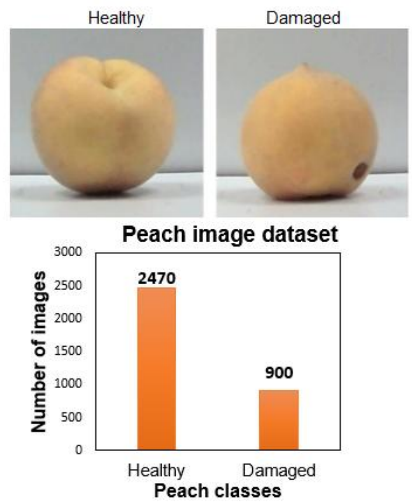

Fig. 2. Number of peach images of each class.

- **Data augmentation**

Data augmentation is a technique used to generate new images from existing raw data by applying various mathematical transformations. This process significantly expands the overall dataset, which in turn improves the training of CNN models by providing more diverse and robust training examples. In this research, data augmentation was implemented to enhance the dataset. Specifically, the RandomHorizontalFlip module from the PyTorch library was utilized. This module randomly flips raw images horizontally, introducing variability and further diversifying the dataset. By applying this transformation, additional images were generated that help the model generalize better and reduce the risk of overfitting. The RandomHorizontalFlip transformation is particularly useful in scenarios where the orientation of objects in the images may vary, such as in the case with peaches. This technique ensures that the model learns to recognize the features of the peaches regardless of their orientation, thereby improving its performance on unseen data.

2.4 Training parameters

Hyperparameters are crucial and adjustable components in CNN models, significantly influencing the training process and performance on specific tasks. The appropriate selection of these hyperparameters is essential to maximize the effectiveness of a CNN model. Among the relevant hyperparameters, optimization algorithms play a critical role in achieving optimal performance.

Stochastic Gradient Descent (SGD) is a widely used optimization algorithm in machine learning tasks, including fruit classification (Elaraby et al., 2022; P & Professor, 2024; Rawung et al., 2023; Singh et al., 2023), and will be employed in this research. SGD updates the model parameters by calculating the gradient of the loss function with respect to a single data point, providing a balance between speed and stability.

One of the key hyperparameters is the “Learning rate”, which controls the speed at which the model adjusts the weight based on the gradient of the loss function (Lopez-Betancur et al., 2024). Specifically, it determines the extent to which the model predictions are modified during each update, based on the error calculated. A high learning rate results in rapid changes to the model parameters, while a low learning rate leads to slower adjustments. Selecting an appropriate learning rate is crucial to ensure that the error decreases effectively and the model converges to the minimum error in the fewest number of epochs (Robles-Guerrero et al., 2024). The “Epoch” hyperparameter indicates the number of times the entire training dataset is passed through the CNN model during the training process. Each epoch constitutes one complete training cycle.

Additionally, the “Batch size” is an important hyperparameter. Models are trained using batches of data, and the batch size specifies the number of data samples processed before updating the model parameters (Lin, 2022). Furthermore, the “Momentum” hyperparameter is employed to accelerate the gradient descent process. Momentum incorporates a fraction of the previous gradient updates when adjusting the current parameters, which helps in speeding up convergence and reducing oscillations. “Weight decay” is a key hyperparameter in DL that helps prevent overfitting by applying a penalty to large weight values. This penalty, incorporated into the loss function, nudges the model towards learning simpler patterns, ensuring it generalizes well to unseen data. Through this regularization technique, DL models maintain robustness and achieve better performance on real-world tasks. The “Folds” are used to split the dataset during cross-validation, allowing the model to be trained and evaluated on different subsets, ensuring a more accurate assessment of its performance. On the other hand, the “Seed” is an initial value used to set up random number generators. In the case of CNNs with TL, setting a seed ensures the reproducibility of results by maintaining consistency in data splitting and weight initialization (Gulzar et al., 2020).

The selection of hyperparameters for this research was conducted in an unbiased manner, without favoring any model. The hyperparameters utilized are listed in Table 2.

Table 2. Tuning hyperparameters in the training process

Hyperparameter	Value
Optimization algorithm	SGD
Learning rate	0.0005
Momentum	0.9
Weight decay	0.0001
Bach size	16
Epochs	50
Seed	40
Fold	5

2.5 Performance Evaluation

2.5.1 Cross Validation

To achieve highly accurate results in a CNN model, it is crucial to work with a well-curated dataset. The dataset is divided into subsets, which are used for both training and validation purposes. This division allows us to calculate the average accuracy of the model across multiple iterations. In the context of this research, we generated five distinct folds from the complete dataset. Each fold serves as a unique subset for training and validation, enabling us to systematically evaluate and identify the optimal model for peach classification. Specifically, in each of the five runs, the dataset is split into 80% for training and 20% for validation. This approach not only enhances the robustness of our model but also ensures that the final accuracy metric is a reliable representation of the model performance.

2.5.2 Matrix confusion

To summarize the results of the training process, it is essential to use a confusion matrix. This tool is a table that allows evaluating the performance of a model by providing a detailed view of its predictions. In the confusion matrix, columns represent the true classes, while rows reflect the classes predicted by the model. The elements on the main diagonal indicate the number of correctly classified cases, while the elements off the diagonal represent the misclassified cases. Analyzing the confusion matrix provides us with key information to adjust and select the optimal set of hyperparameters for the CNN model, thus facilitating the improvement of its performance on specific tasks.

In the confusion matrix, there are four relevant terms: True Positives (TP), which indicate the number of correctly classified cases; True Negatives (TN), which refer to the number of items correctly identified as not belonging to the class; False Positives (FP), which are the items incorrectly classified as belonging to the class when in fact they are not; and False Negatives (FN), which are the cases in which the model incorrectly predicts that an item does not belong to the class when it does (Lopez-Betancur et al., 2022). These metrics are essential to calculate model performance and tune the hyperparameters effectively. These terms are essential for evaluating the quality of model predictions and for calculating key performance metrics such as accuracy, precision, recall, specificity, F1-score, G-mean, and Index of Balanced Accuracy (IBA) (Guerrero-Mendez et al., 2020).

Accuracy: This measures how often the model makes correct predictions. It is calculated as:

$$Accuracy = \frac{TP+TN}{(TP+TN+FP+FN)}. \tag{1}$$

Although accuracy gives a general idea of performance, it can be misleading in imbalanced datasets.

Precision: This indicates the precision of positive predictions. It is the ratio of true positives to all positive predictions:

$$Precision = \frac{TP}{TP+FP}. \tag{2}$$

High precision means the model is good at identifying peaches without many false alarms.

Recall (Sensitivity): This measures the model ability to detect all actual positive cases:

$$Recall = \frac{TP}{TP+TN}. \tag{3}$$

High recall indicates that the model finds most of the actual peaches.

Specificity: This assesses the model ability to correctly identify negative cases:

$$Specificity = \frac{TN}{TN+FP}. \tag{4}$$

High specificity means the model effectively identifies non-peaches.

F1-Score: This combines precision and recall into a single metric, providing a balance between them:

$$F1 - score = \frac{2TP}{2TP+FP+FN}. \tag{5}$$

It is useful for understanding the trade-off between precision and recall.

Geometric Mean (G-mean): These balances recall and specificity by taking their square root product:

$$G_{mean} = \sqrt{\frac{TP}{TP+FN} \times \frac{TN}{TN+FP}}\,. \tag{6}$$

It is particularly useful for imbalanced datasets.

Index of Balanced Accuracy (IBA): This weight measure combines sensitivity and specificity, using a weighting factor of 0.1 by (Maeda-Gutiérrez et al., 2020; Navarro-Solís et al., 2024):

$$IBA = \left(1 + 0.1\left(\frac{TP}{TP+FN} - \frac{TN}{TN+FN}\right)\right)\left(\frac{TP}{TP+FN} \times \frac{TN}{TN+FP}\right). \tag{7}$$

IBA provides a balanced evaluation of the model performance across all classes.

These metrics collectively offer a comprehensive view of the model performance, especially in scenarios with imbalanced datasets like peach classification. While accuracy provides an overall picture, precision, recall, specificity, F1-score, G-mean, and IBA offer deeper insights into the model ability to handle different aspects of the classification task.

The computer specifications for the training process were as follows: an 11th Gen Intel® Core™ i7-11700K processor, 32GB of RAM, an NVIDIA RTX 3060 12GB graphics card, and the Windows 11 Pro operating system. The implemented algorithms were executed using Python 3, an open-source programming language. Furthermore, the CNN models were trained using PyTorch library, specifically version 1.9.1. In addition, to optimize the training and evaluation process, we leveraged the Torchvision package, which provides a vast array of pre-trained models and is indispensable for developing advanced computer vision applications.

3 Results and Discussion

For the classification of peaches, a cross-validation process was conducted using five folds for each CNN model. The validation score of each fold was utilized to compute the mean validation score for each model. Based on these results, the optimal model for classifying peaches was identified. Specifically, in each fold, 80% of the image dataset was allocated for training, while the remaining 20% was reserved for validation. Additionally, it is important to note that the images used in each fold are different from those in the other folds, meaning that the images rotate and are distributed differently in each iteration. This approach ensured that the models were exposed to a diverse range of variations, thereby improving their generalization capabilities.

The results obtained in this study demonstrate the effectiveness of ResNet models for the classification of healthy and damaged peaches. The average performance of the models exceeded 95%, evidencing their ability to discriminate between both classes. Table 3 presents the validation accuracy for each fold and the mean accuracy for each ResNet model. ResNet-50 emerged as the top-performing model with a mean accuracy of 95.96%, outperforming other models such as ResNet-18, ResNet-34, ResNet-101, and ResNet-152. Notably, ResNet-50 achieved the highest accuracy in Fold 4 (96.44%) and consistently performed well across all folds, indicating its robust generalization.

Table 3. Validation accuracy (%) in each fold for every CNN model

CNN model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean Fold
ResNet-18	94.36	96.29	94.96	96.29	95.99	95.58
ResNet-34	95.40	96.14	94.36	96.14	96.28	95.66
ResNet-50	95.85	96.29	94.81	96.44	96.44	95.96
ResNet-101	93.91	96.29	95.10	96.88	96.30	95.69
ResNet-152	95.25	96.299	94.51	95.85	96.44	95.67

ResNet-152, despite its depth, did not outperform ResNet-50, which suggests that deeper models may not always yield better performance, especially when the dataset is not sufficiently large or complex to benefit from the additional layers. This observation aligns with a previous study that have noted a potential trade-off between model depth and performance in certain tasks (He et al., 2016).

For a more detailed analysis of the learning process of the five architectures, it is possible to examine the confusion matrices obtained for each of the selected models. Figure 3 presents the confusion matrices for each model, which reflect the results of the best epoch in each fold.

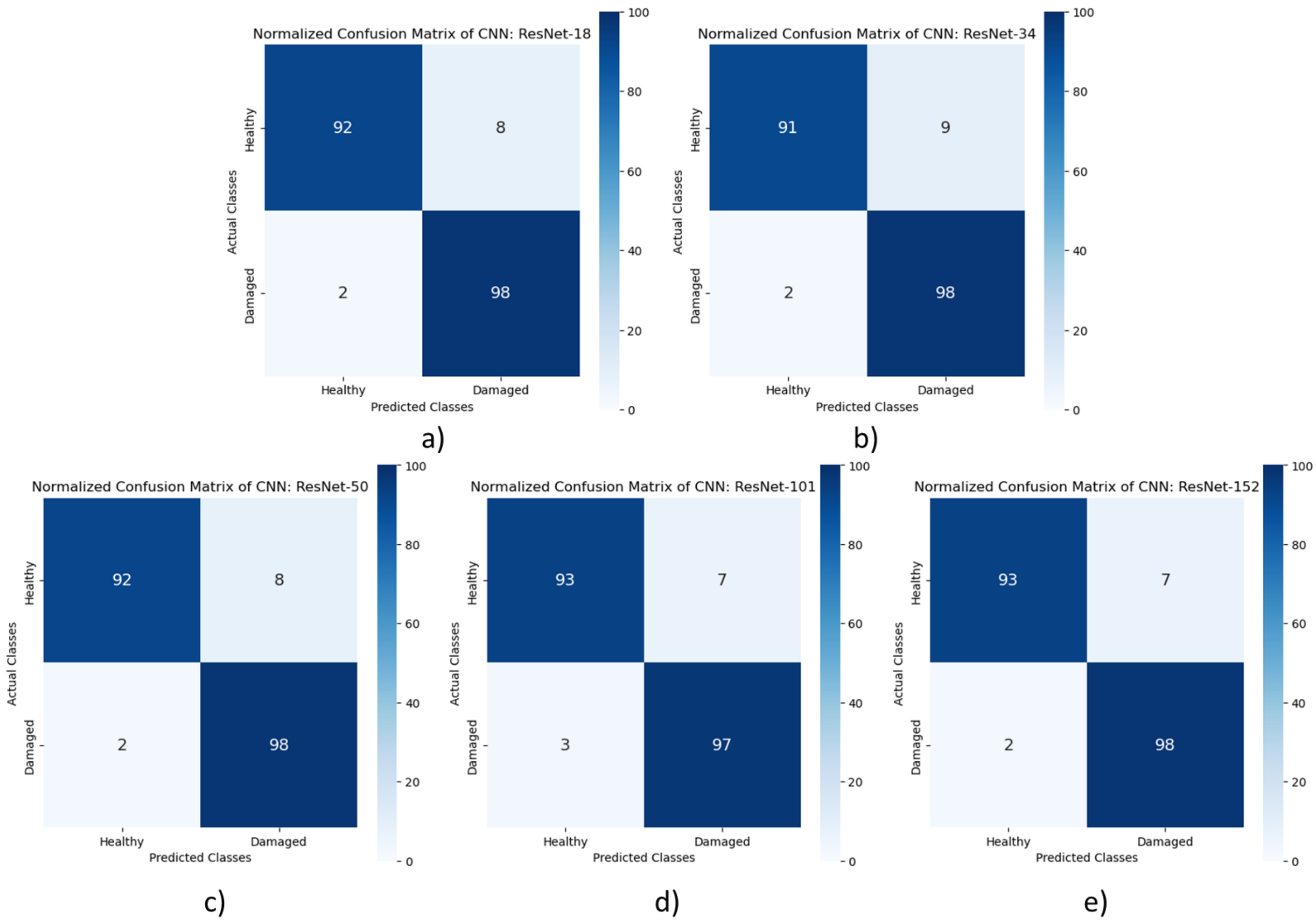

Fig. 3. Confusion matrices of the architectures: a) ResNet-18, b) ResNet-34, c) ResNet-50, d) ResNet-101 and e) ResNet-152.

The confusion matrices reveal that all the models achieved remarkably high performance, demonstrating their ability to effectively distinguish between the two classes. The high values along the main diagonal of each confusion matrix highlight the overall accuracy of the models. However, subtle variations in performance were observed among the different architectures, suggesting that the depth and architectural complexity of the networks influence their classification capabilities.

Among the evaluated models, ResNet-50 stood out for its balance between accuracy and efficiency. Specifically, ResNet-50 correctly identified 92% of healthy peaches and 98% of damaged peaches. Despite its strong performance, the model exhibited an 8% false positive rate, where some healthy peaches were misclassified as damaged, and a 2% false negative rate, where damaged peaches were incorrectly classified as healthy.

When compared to ResNet-18 and ResNet-34, ResNet-50 demonstrated a slight advantage, particularly in the classification of damaged peaches. Its performance was comparable to that of deeper models like ResNet-101 and ResNet-152, but with lower computational requirements. These results suggest that ResNet-50 achieves an optimal balance between depth and performance, making it a versatile option for this specific classification task.

All these pre-trained models were evaluated using various metrics: accuracy, precision, sensitivity, specificity, F-Score, G_mean and IBA, using the same hyperparameters. Table 4 provides the mean performance metrics for each model, including precision, recall, specificity, F1-score, G-mean, and IBA. ResNet-50 consistently outperformed the other models in most of these metrics. For instance, ResNet-50 achieved a precision of 95.94%, a recall of 95.96%, and an F1-score of 95.99%, all of which are superior to the other models. These results indicate that ResNet-50 not only has high accuracy but also effectively balances precision and recall, making it a robust choice for peach classification.

The specificity of ResNet-50 (92.16%) suggests that it is effective in correctly identifying healthy peaches, while its G-mean (93.97%) and IBA (88.64%) further confirm its balanced performance across both classes. These metrics are crucial in ensuring that the model does not exhibit bias towards one class, which is particularly important in applications where both classes are equally important.

Table 4. Mean scores of performance metrics (%)

CNN model	Precision	Recall	Specificity	FI-score	G_mean	IBA
ResNet-18	95.57	95.57	90.88	95.52	93.08	87.05
ResNet-34	95.66	95.66	92.56	95.65	94.05	88.73
ResNet-50	95.94	95.96	92.16	95.99	93.97	88.64
ResNet-101	95.71	95.69	91.17	95.63	93.27	87.38
ResNet-152	95.62	95.67	92.32	95.64	93.91	88.50

Table 5 presents the training time for each model across the five folds. As expected, deeper models such as ResNet-101 and ResNet-152 required more time to train compared to shallower models like ResNet-18 and ResNet-34. ResNet-152, with its 152 layers, took approximately 40.65 minutes per fold, which is significantly longer than ResNet-50, which took about 20.23 minutes per fold. Despite the longer training time, ResNet-50 performance justifies its training time, especially considering its superior accuracy and generalization. This trade-off between training time and performance is a common consideration in DL, where deeper models often require more computational resources but may not always provide commensurate improvements in performance (LeCun et al., 2015).

Table 5. Training time (min)

CNN model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean Fold
ResNet-18	11.48	11.27	11.27	11.3	11.28	11.32
ResNet-34	15.37	15.42	15.37	15.38	15.4	15.39
ResNet-50	20.28	20.1	20.05	20.17	20.57	20.23
ResNet-101	30.4	30.5	30.53	30.53	30.4	30.47
ResNet-152	40.75	40.68	40.62	40.62	40.6	40.65

The performance of ResNet-152, despite its depth, was not superior to ResNet-50, which may be attributed to the complexity of the dataset or the potential for overfitting in deeper models. This finding highlights the importance of selecting a model that is appropriately sized for the task at hand, rather than deeper architectures.

The cross-validation approach used in this study ensured that the models were evaluated on multiple subsets of the data, providing a more reliable estimate of their generalization. The consistent performance of ResNet-50 across all folds further reinforces its suitability for this application.

In summary, this study provides evidence that ResNet-50 is an effective model for classifying healthy and damaged peaches. Its performance metrics and training time make it a practical choice for implementation in agricultural settings, where efficiency and accuracy are paramount.

4 Conclusions

This study provides compelling evidence that ResNet models, particularly ResNet-50, are highly effective for the classification of healthy and damaged peaches. The superior performance of ResNet-50, coupled with its computational efficiency, makes it a practical choice for implementation in agricultural settings. The results underscore the potential of DL techniques to enhance quality control processes in the agri-food industry, thereby reducing post-harvest losses and ensuring the delivery of high-quality produce to consumers.

The findings of this research have significant implications for the agricultural sector. By automating the classification process, peach producers can achieve higher efficiency, reduce labor costs, and improve the overall quality of their products. Furthermore, the use of ResNet-50 in peach classification can be extended to other fruit varieties, offering a scalable solution for various agricultural applications. In summary, this study demonstrates the potential of ResNet-50 as a powerful tool for peach classification, paving the way for its adoption in the agricultural industry to improve quality control and reduce post-harvest losses.

References

Africano P., K. L., Almanza-Merchán, P. J., & Balaguera-López, H. E. (2015). Fisiología y bioquímica de la maduración del fruto de durazno [Prunus persica(L.) Batsch]. Una Revisión. *Revista Colombiana de Ciencias Hortícolas*, *9*(1), 161–172. https://doi.org/10.17584/rcch.2015v9i1.3754

Akbar, M., Ullah, M., Shah, B., Khan, R. U., Hussain, T., Ali, F., Alenezi, F., Syed, I., & Kwak, K. S. (2022). An effective deep learning approach for the classification of Bacteriosis in peach leave. *Frontiers in Plant Science*, *13*. https://doi.org/10.3389/fpls.2022.1064854

Alebiosu, D. O., & Muhammad, F. P. (2019). Medical Image Classification: A Comparison of Deep Pre-trained Neural Networks. *2019 IEEE Student Conference on Research and Development (SCOReD)*, 306–310. https://doi.org/10.1109/SCORED.2019.8896277

Corvalán, J. G. (2018). Inteligencia artificial: Retos, desafíos y oportunidades - Prometea: la primera inteligencia artificial de Latinoamérica al servicio de la Justicia. *Revista de Investigações Constitucionais*, *5*, 295–316. https://doi.org/10.5380/rinc.v5i1.55334

Diario Oficial de la Federacion. DECLARATORIA de vigencia de la Norma Mexicana NMX-FF-060-SCFI-2009. (n.d.). Retrieved September 17,2024, from https://www.dof.gob.mx/

Elaraby, A., Hamdy, W., & Alanazi, S. (2022). Classification of Citrus Diseases Using Optimization Deep Learning Approach. *Computational Intelligence and Neuroscience*, *2022*(1), 9153207. https://doi.org/10.1155/2022/9153207

Flores, F. A. I., Sanchez, D. L. C., Urbina, R. O. E., Coral, M. Á. V., Medrano, S. E. V., & Gonzales, D. G. E. (2022). Inteligencia artificial en educación: Una revisión de la literatura en revistas científicas internacionales. *Apuntes Universitarios*, *12*(1), Article 1. https://doi.org/10.17162/au.v12i1.974

Gonzales, C. E. S., Zapata, M. S. C., & Zapata, E. Y. C. (2022). Determinación del estado de madurez de durazno crema huayco mediante procesamiento de imágenes con Raspberry Pi. *INGnosis*, *8*(2), Article 2. https://doi.org/10.18050/ingnosis.v8i2.2830

Guerrero-Mendez, C., Saucedo-Anaya, T., Moreno, I., Araiza-Esquivel, M., Olvera-Olvera, C., & Lopez-Betancur, D. (2020). Digital Holographic Interferometry without Phase Unwrapping by a Convolutional Neural Network for Concentration Measurements in Liquid Samples. *Applied Sciences*, *10*(14), Article 14. https://doi.org/10.3390/app10144974

Gulzar, Y., Hamid, Y., Soomro, A. B., Alwan, A. A., & Journaux, L. (2020). A Convolution Neural Network-Based Seed Classification System. *Symmetry*, *12*(12), Article 12. https://doi.org/10.3390/sym12122018

He, K., Zhang, X., Ren, S., & Sun, J. (2016). *Deep Residual Learning for Image Recognition*. 770–778. https://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Lin, R. (2022). Analysis on the Selection of the Appropriate Batch Size in CNN Neural Network. *2022 International Conference on Machine Learning and Knowledge Engineering (MLKE)*, 106–109. https://doi.org/10.1109/MLKE55170.2022.00026

Lopez-Betancur, D., González-Ramírez, E., Guerrero-Mendez, C., Saucedo-Anaya, T., Rivera, M. M., Olmos-Trujillo, E., & Gomez Jimenez, S. (2024). Evaluation of Optimization Algorithms for Measurement of Suspended Solids. *Water*, *16*(13), Article 13. https://doi.org/10.3390/w16131761

Lopez-Betancur, D., Moreno, I., Guerrero-Mendez, C., Saucedo-Anaya, T., González, E., Bautista-Capetillo, C., & González-Trinidad, J. (2022). Convolutional Neural Network for Measurement of Suspended Solids and Turbidity. *Applied Sciences*, *12*(12), Article 12. https://doi.org/10.3390/app12126079

Lopez-Betancur, D., Saucedo-Anaya, T., Guerrero-Mendez, C., Navarro-Solís, D., Silva-Acosta, L., Robles-Guerrero, A., & Gomez-Jimenez, S. (2024). Evaluating CNN Models and Optimization Techniques for Quality Classification of Dried Chili Peppers (Capsicum annuum L.). *International Journal of Combinatorial Optimization Problems and Informatics*, *15*(2), Article 2. https://doi.org/10.61467/2007.1558.2024.v15i2.462

Maeda-Gutiérrez, V., Galván-Tejada, C. E., Zanella-Calzada, L. A., Celaya-Padilla, J. M., Galván-Tejada, J. I., Gamboa-Rosales, H., Luna-García, H., Magallanes-Quintanar, R., Guerrero Méndez, C. A., & Olvera-Olvera, C. A. (2020). Comparison of Convolutional Neural Network Architectures for Classification of Tomato Plant Diseases. *Applied Sciences*, *10*(4), Article 4. https://doi.org/10.3390/app10041245

Medina, G., Chuk, O., Luna, A., & Bertero, R. (2022). *Redes neuronales convolucionales para determinar redondez en partículas de arena*.

Méndez Almansa, J. E., & Silva Salamanca, J. S. (2022). *Desarrollo de una aplicación móvil para el reconocimiento de la madurez de un grupo de frutas a través del análisis de imágenes por medio de redes neuronales*. http://repository.unipiloto.edu.co/handle/20.500.12277/11498

Naranjo-Torres, J., Mora, M., Hernández-García, R., Barrientos, R. J., Fredes, C., & Valenzuela, A. (2020). A Review of Convolutional Neural Network Applied to Fruit Image Processing. *Applied Sciences*, *10*(10), Article 10. https://doi.org/10.3390/app10103443

Navarro-Solís, D., Guerrero-Méndez, C., Saucedo-Anaya, T., Lopez-Betancur, D., Silva, L., Robles-Guerrero, A., & Gómez-Jiménez, S. (2024). Analysis of Convolutional Neural Network Models for Classifying the Quality of Dried Chili Peppers (Capsicum Annuum L). En H. Calvo, L. Martínez-Villaseñor, H. Ponce, R. Zatarain Cabada, M. Montes Rivera, & E. Mezura-Montes (Eds.), *Advances in Computational Intelligence. MICAI 2023 International Workshops* (pp. 116–131). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-51940-6_10

P, P., & Professor, B. B. A. (2024). A Comparison of Different Optimization Techniques in AlexNet for Plant Disease Classification: SGD, Adadelta, ASGD and Adagrad. *2024 5th International Conference on Smart Electronics and Communication (ICOSEC)*, 1907–1914. https://doi.org/10.1109/ICOSEC61587.2024.10722248

Rawung, B. H., Djamal, E. C., & Yuniarti, R. (2023). Classification of lemon fruit ripe using convolutional network. *AIP Conference Proceedings*, *2714*(1), 030029. https://doi.org/10.1063/5.0129395

Robles-Guerrero, A., Gómez-Jiménez, S., Saucedo-Anaya, T., López-Betancur, D., Navarro-Solís, D., & Guerrero-Méndez, C. (2024). Convolutional Neural Networks for Real Time Classification of Beehive Acoustic Patterns on Constrained Devices. *Sensors*, *24*(19), Article 19. https://doi.org/10.3390/s24196384

Singh, R., Sharma, N., & Gupta, R. (2023). Impact of Different Optimizers on InceptionResNet V2 for Amazon Fruit Classification. *2023 International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE)*, 1–6. https://doi.org/10.1109/RMKMATE59243.2023.10369018

Taye, M. M. (2023). Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions. *Computation*, *11*(3), Article 3. https://doi.org/10.3390/computation11030052

Yao, N., Ni, F., Wu, M., Wang, H., Li, G., & Sung, W.-K. (2022). Deep Learning-Based Segmentation of Peach Diseases Using Convolutional Neural Network. *Frontiers in Plant Science*, *13*. https://doi.org/10.3389/fpls.2022.876357