Ingeniería Investigación y Tecnología
Volume XXVI (Issue 3), July-September 2025 1-12
ISSN 2594-0732 FI-UNAM Peer-reviewed article
Article information: Received: July 9, 2023; accepted: April 29, 2025

https://doi.org/10.22201/fi.25940732e.2025.26.3.019

Ingeniería Investigación y Tecnología

Deep learning to identifying food maturity for prevent food waste

Aprendizaje profundo para identificar la madurez de los alimentos y prevenir el desperdicio de alimentos

Cuaya-Simbro German (Corresponding author)
Tecnológico Nacional de México
ITS del Oriente del Estado de Hidalgo
E-mail: gcuaya@itesa.edu.mx
https://orcid.org/0000-0001-6303-154X

Cuevas-Ordoñez David Williams
Tecnológico Nacional de México
ITS del Oriente del Estado de Hidalgo
E-mail: 16030583@itesa.edu.mx
https://orcid.org/0009-0009-7680-8296

Abstract

The article discusses the issues of malnutrition and food insecurity worldwide, highlighting the significant problem of food waste. We propose a technological solution to address this specific issue. The use of computer vision, specifically Deep Learning (DL) models, is proposed to accurately determine food maturity stages, thereby assessing their suitability for consumption and preventing unnecessary waste. This manuscript explores various applications of deep learning in topics related to the food industry, focusing on research aimed at mitigating food waste. Three datasets were constructed to identify the most commonly wasted foods, as well as to determine the maturity levels of these foods. A deep learning model was developed to identify three different food items with an effectiveness rate of 77 %. Additionally, three deep learning models were created to assess the maturity degree of each food, achieving an average effectiveness of 70 %. These results are comparable to similar studies that have employed deep learning models. This research shows the feasibility of using deep learning models to determine the maturity levels of three types of food that are frequently wasted in México. This approach is novel in research areas dedicated to developing methods, techniques, or technologies for avoiding or reducing food waste.
Keywords: Deep learning, food maturity, prevent food waste, computer vision, food dataset.

Resumen

El artículo aborda los problemas de desnutrición e inseguridad alimentaria en todo el mundo, destacando el importante problema del desperdicio de alimentos. Proponemos una solución tecnológica para abordar este problema específico. Se propone el uso de la visión por computadora, en particular modelos de Aprendizaje Profundo (Deep Learning), como medio para identificar el estado de madurez de los alimentos, determinando así su comestibilidad y previniendo el desperdicio prematuro. Este manuscrito explora diversas aplicaciones del aprendizaje profundo en temas relacionados con la industria alimentaria, centrándose en investigaciones orientadas a mitigar el desperdicio de alimentos. Se construyeron tres conjuntos de datos para identificar los alimentos comúnmente más desperdiciados, así como para determinar los niveles de madurez de estos alimentos. Se desarrolló un modelo de aprendizaje profundo para identificar tres alimentos diferentes con una tasa de efectividad de 77 %. Además, se crearon tres modelos de aprendizaje profundo para evaluar el grado de madurez de cada alimento, logrando una efectividad promedio de 70 %. Estos resultados son comparables a estudios similares que han utilizado modelos de aprendizaje profundo. Esta investigación muestra la viabilidad de utilizar modelos de aprendizaje profundo para determinar los niveles de madurez de tres tipos de alimentos que se desperdician con frecuencia en México. Este enfoque es novedoso en áreas de investigación dedicadas al desarrollo de métodos, técnicas o tecnologías para evitar o reducir el desperdicio de alimentos.
Descriptores: Aprendizaje profundo, madurez de alimentos, prevención de desperdicio de alimentos, visión por computadora, conjunto de datos de alimentos.

Introduction

In the context of increasing malnutrition and food insecurity in Africa and America, malnutrition is recognized as a global pandemic (Berruguete, 2019). Food waste has emerged as a major obstacle preventing access to food for certain segments of the population. It occurs at the market and consumer levels, when food remains unsold or untouched in households. Food waste not only contributes to 8 % of global greenhouse gas emissions but also wastes 30 % of agricultural land and 21 % of fresh water resources, as reported by the FAO (UN, 2016). Thus, reducing food waste is crucial.

Different researches such as Moreno (2016), Kunszabó *et al.* (2019), and Mozos *et al.* (2020) describe the problem of food loss and waste and its relationship with food security, care for the environment, and natural resources. Such research proposes solutions and possible lines of action so that health, nutrition, and food professionals, as well as other professionals, can contribute to making better use of food in society.

One alternative to support the effort to reduce food waste is the use of computational technologies, specifically computer vision, to identify the maturity state of the food. In recent years, there has been a significant advance in this technology with the integration of machine learning techniques, particularly Deep Learning (DL) models. The use of DL models in computer vision has improved the efficiency of this technology and has been successfully applied in different areas such as medicine, security, biology, traffic, and agriculture (Balas, 2019; Patel & Thakkar, 2020; Chai *et al.*, 2021; Li, 2022). Additionally, there are researchers that have reported the application of machine learning (ML) techniques to support activities related to the food industry, such as food production improving methods of crop farming, cultivation, and processing (Kakani *et al.*, 2020; Kumar *et al.*, 2021).

Several authors have reported that Artificial Neural Networks (ANN) models are widely applied in the food industry for prediction, classification, and control tasks, as well as for food processing and technology. The supervised ANN method has the ability to learn from examples, which allows for the prediction process to be done accurately. Meanwhile, the unsupervised method of ANN is more commonly found for the classification task (Liu *et al.*, 2019; Koszela *et al.*, 2017; Chasiotis *et al.*, 2020; Dang *et al.*, 2019).

Although previous research has generally focused on enhancing food production processes, there remains a limited exploration of computational strategies specifically aimed at preventing food waste, which constitutes a clear research opportunity.

Food waste management has captured the attention of environmental associations, policy makers, analysts, and researchers (Morone *et al.*, 2019), due to its social, eco nomic, and environmental impacts (Carvajal *et al.*, 2021). The generation of food waste at both the supplier and consumer levels stems from a complex set of inter acting behaviors. Computational and mathematical models provide various methods to simulate, diagnose, and predict different aspects within the complex system of food waste generation and prevention (Reynolds, 2020).

Related with computational models, is the use of machine learning techniques to mitigate food waste is. Among the research related to this topic, we can cite the following:

In the study (Panda & Dwivedi, 2020), machine learning techniques were used to address food waste in educational institutions. They developed classification models to predict food usage, comparing decision tree and naive Bayes algorithms achieved an average accuracy of 0.74, while decision tree achieved 0.65. Their focus was on determining if individuals would use or waste food. In contrast, our proposal aims to determine the maturity of food and recommend appropriate actions to prevent spoilage or rotting. Similarly, in the research conducted by Zhang *et al.* (2021), computer vision was employed to classify and sort domestic waste based on regulations. They introduced the Waste Recognition-Retrieval (W2R) algorithm, consisting of two stages: a Recognition Model (RegM) and a Recognition-Retrieval Model (RevM). Results showed that the average accuracy of RevM (94.71 % ± 1.69) was significantly higher than that of ClfM-VGG (69.66 % ± 3.43) and MS (72.50 % ± 11.37). Their study, like the mentioned thesis, focused on identifying various types of waste, including food waste, to determine recyclability. The results contribute to reducing food waste and assisting end-users in distinguishing between waste and valuable items.

In Farinella *et al.* (2020), a system was developed to measure food waste using a camera and a database of various food types. Key factors considered were the initial food quantity and the amount wasted after disposal. Analyzing these data values allowed for improving production efficiency and reducing future food waste. Machine learning techniques were employed, specifically training a convolutional neural network, to identify different food items. Although the authors trained the system to recognize 10 common dining hall food items, they did not report the accuracy of the CNN model. While an object detection model was constructed for food, the authors did not address the maturity aspect of the food items.

In Kakani *et al.* (2020), an autonomous warehouse system leveraging Machine Learning and Blockchain technology was proposed to address food wastage. The system consisted of three interconnected subsystems aimed at reducing food waste. The first subsystem focused on predicting food product expiry dates based on current warehouse environmental conditions. While their research shares similarities with our objective, they utilized an advanced AI model to predict the remaining shelf life of products under current conditions. In contrast, our model assesses food maturity and estimates the time before spoilage using visual degradation indicators. This approach differs from expiry dates, as some food can still be consumed beyond that point, helping to mitigate significant food waste. The authors implemented and evaluated their model using the Egg Plant (Brinjal) dataset, incorporating real-time environmental features such as temperature, humidity, ice coverage, water sprinkling, and carbon dioxide levels. They achieved a minimum training error of 0.4268 by minimizing the model's mean squared error using the RMSprop optimizer.

In Kumar *et al.* (2022) , the authors use images to identify healthy and deteriorated fruits, including apples, bananas, and oranges. For classifying photographs into fresh and decaying fruits, softmax is used, while CNN obtains fruit image properties. They use a dataset from Kaggle to evaluate the suggested model's performance, and it achieves a 97.14 percent accuracy rate. This research is similar to our research because the authors' model identifies the physical condition of a food. It is possible to assume that the deteriorate class corresponds to a rotten food, and then it is possible to suggest some action to take with the food before it reaches that physical state. In contrast, the model that we built considers different physical conditions or maturity levels, i.e., multi-class classification, which is a more complex problem.

Finally, in Bhargava & Bansal (2021), a critical comparison of different algorithms proposed by researchers for quality inspection of specific types of food has been carried out. However, our research differs from this study, as we are focused on identifying the maturity, not the quality, of a specific food using a deep learning model.

In summary, various studies have focused on reducing food waste using computational techniques, including computer vision and deep learning models. While some studies identify food waste and suggest preventive actions, our research specifically focuses on using deep learning models to identify food waste based on visual attributes. Our contributions include: developing datasets for three types of food (potato, tomato, and mango) to determine their maturity levels, creating an object identification model capable of distinguishing between these foods, and assessing their expiration status.

Materials and methods

Concepts related with food maturity

Food's labeling

In many foods that people consume, there are indications related to the maturity condition of the food, such as food labeling. The labeling influences consumers to make well-informed decisions. It is important to clarify that the indication "use before" or "expiration date" refers to safety, whereas the label "use preferably before" pertains to quality considerations. It is important to distinguish between the expiration date and preferential consumption. The first indicates highly perishable products with a microbiological risk. Once exceeded, the product should not be consumed, as there is a risk that it is in poor condition and contains pathogenic bacteria. The second applies to durable products, which may lose some of their properties once the indicated date has elapsed, but they do not present a microbiological risk. Unfortunately, not all food has this labeling, specifically fruit or vegetable products.

Food maturity

Post-harvest physiologists distinguish three stages in the lifespan of fruits and vegetables: maturation, ripening, and senescence. Maturation denotes that the edible portion of the fruit or vegetable has reached its full size, although it might not be ready for immediate consumption. Ripening occurs after or during maturation, making the produce edible, as evidenced by its taste. Senescence is the last stage, characterized by the natural degradation of the fruit or vegetable, resulting in the loss of texture, flavor, etc. Senescence ends with the death of the tissue of the fruit, but until this point, the food is consumable, despite its aesthetic condition. It is important to note that during the senescence stage, the food is still safe to eat, even if it has reached its expiration date. However, customers often tend to waste the food when they perceive it as unappealing, despite it being still edible.

Food maturity index

Some typical maturity indexes are described in Hathi *et al.* (2020), but we used a visual index, based on color, which establishes that the loss of green color in many fruits is a valuable guide to maturity. We used this index to assign a maturity level (class) to mangoes, potatoes, and tomatoes.

Machine learning

Machine learning (ML) is a key component of artificial intelligence (AI) that utilizes data and algorithms to mimic human learning and improve accuracy. ML classifiers are widely used in AI applications, enabling automated data analysis, process optimization, and valuable insights extraction.

In ML, there are two primary types of classification algorithms:

Binary classification: This supervised learning task involves categorizing data into two distinct and mutually exclusive groups or categories. Labels such as 0 and 1, positive and negative, true or false are assigned to these groups. Binary classification models are trained using labeled datasets, where each data point is associated with the desired outcome.

Multiclass classification: This type of supervised learning problem aims to classify data into three or more categories. Unlike binary classification, which predicts between two classes, multiclass classifiers are trained to predict from three or more classes.

Deep learning

Concept

It is a machine learning technique that teaches computers to learn by example, just as humans do. In deep learning, a computer model learns to perform classification tasks directly from images, text, or sound. Deep learning models can achieve high accuracy, sometimes even exceeding human-level performance. These models are trained using a large labeled dataset and neural network architectures that contain many layers (Umeki, 2014).

The majority of deep learning approaches employ neural network structures, which is why deep learning models are frequently referred to as deep neural networks. The term 'deep' typically pertains to the quantity of hidden layers within the neural network. Within artificial neural networks, the convolutional neural network (ConvNets or CNN) is one of the main categories for doing image recognition, image classification, object detection, and more.

RetinaNet architecture

RetinaNet is a unified network that consists of a backbone network and two task specific subnetworks, illustrated in Figure 1. The backbone network is responsible for generating a convolutional feature map across the entire input image and is typically a pre-existing convolutional network. The first subnetwork performs convolutional object classification on the output of the backbone, while the second subnetwork handles convolutional bounding box regression (Lin *et al.*, 2017).

Assessment measures for machine learning models

Confusion matrix

A confusion matrix describes the performance of a classification model (Barrios, 2019). Put simply, it provides a summary of the model's performance. Each column in the matrix corresponds to the number of predictions made for each class, while each row represents the instances belonging to the actual class. The confusion matrix yields the following computed values:

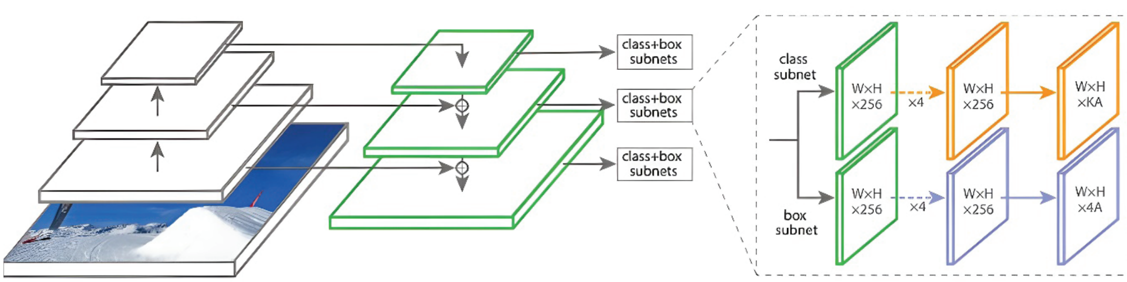

Figure 1. The RetinaNet architecture (Lin *et al.*, 2017) employs the following components: a) a backbone network that generates a comprehensive, multi-scale convolutional feature pyramid b). To this backbone, RetinaNet adds two subnetworks: one for classifying anchor boxes c) and another for regressing from anchor boxes to ground-truth object boxes d). Additionally, RetinaNet incorporates a focal loss function that addresses the accuracy discrepancy between one-stage detectors and traditional two-stage detectors

TN (True Negative): This refers to the number of outcomes that were originally negative and correctly predicted as negative.
FP (False Positive): This represents the count of outcomes that were originally negative but were falsely predicted as positive. This error is also known as a type 1 error.
FN (False Negative): This indicates the number of outcomes that were originally positive but were mistakenly predicted as negative. This error is also referred to as a type 2 error.
TP (True Positive): This denotes the number of outcomes that were originally positive and correctly predicted as positive. The objective of machine learning and deep learning models is to maximize TN and TP while minimizing FN and FP.

Binary model performance measures

The following measures are used to assess binary models:

Precision: This metric can be used to measure the overall effectiveness of a deep learning model in identifying or detecting a class. The formula to obtain its value is:

$$\frac{TP}{TP+FP}$$

Recall: This metric tells us how many of the classes that were identified within an image have been correctly identified. The formula to obtain its value is:

$$\frac{TP}{TP+FN}$$

F1 Score: The F1 score integrates precision and recall measures into a single value, providing an indication of the deep learning model's effectiveness in distinguishing and detecting different classes. It offers insight into the model's ability to accurately identify and differentiate between various categories. The formula used to compute the F1 score is:

$$2*\frac{Precision*Rcall}{Precision+Rcall}$$

Accuracy: This metric measures the percentage of cases where the model is correct. The formula to calculate its value is:

$$\frac{TP+TN}{TP+TN+FP+FN}$$

Multi-class model performance measures

When a model has more than two classes, like in our case, the previously mentioned metrics are not sufficient to assess the performance of the model. Therefore, it is necessary to compute the following values:

Micro Average: The micro-average focuses on analyzing individual classes. Micro average precision and recall scores are derived from the true positives (TPs), true negatives (TNs), false positives (FPs), and false negatives (FNs) of each class within the model:

$$microavgPrecision=\frac{TP_1+TP_2+...+TP_n}{(TP_1+TP_2+...+TP_n)+(FP_1+FP_2+...+FP_n)}$$

$$microavgRecall=\frac{TP_1+TP_2+...+TP_n}{(TP_1+TP_2+...+TP_n)+(FN_1+FN_2+...+FN_n)}$$

micro avg F1 *Score = micro avg Recall*

Macro Average: This metric is an average of Precision, Recall, and F1 Score. It is used when all classes should be treated equally to assess the overall performance of the classifier against the most frequent class labels. To obtain its value, the arithmetic mean of the Precision and Recall scores is calculated, and the F1 Score is calculated using the harmonic mean of the previous two:

$$macro\,avg\,Precision=\frac{Prec_1+Prec_2+...+Prec_n}{n}$$

$$macro\,avg\,Recall=\frac{Recall_1+Recall_2+...+Recall_n}{n}$$

$$macro\,avg\,\mathrm{F1}\,Score=\frac{\mathrm{F1}-score_1+\mathrm{F1}-score_2+...+\mathrm{F1}-score_n}{n}$$

Weighted Average: Computes metrics for each class and finds its support-weighted average (the number of true instances for each class). It is used when there are un balanced classes, and it is also calculated based on the Precision per class, like macro, but it takes into account the number of samples of each class in the data. The formula to calculate its value is the following, where P is Precision or Recall, N is Number of samples, and C is class:

$$Weightedavg=\frac{P_{C1}*N_{C1}+P_{C2}*N_{C2}+P_{C3}*N_{C3}}{TotalNumberofsamples}$$

https://doi.org/10.22201/fi.25940732e.2025.26.3.019

Programming environment

Google Colab

Google Colab is a cloud-based service that is based on Jupyter Notebooks. It allows users to use Google's GPUs and TPUs for free and comes with a range of libraries including Scikit-learn, PyTorch, TensorFlow, Keras, and OpenCV. It is an ideal tool for practicing and improving knowledge of Data Science techniques and tools, as well as for developing Machine Learning and Deep Learning applications.

Python

Python is commonly employed as an interpreted language, eliminating the need for compiling or linking and speeding up program development. It has emerged as one of the primary programming languages for Artificial Intelligence applications.

TensorFlow

TensorFlow is an open-source software library for numerical computing using data flow graphs. The nodes on the graphs represent mathematical operations, while the edges represent the multidimensional data arrays (tensors) communicating with each other. TensorFlow is a great platform for building and training neural networks, which can detect and identify patterns and correlations, similar to how humans learn and reason (Tensorflow, 2019).

TensorFlow is Google's open-source platform for machine learning, and is currently the most widely used tool for building deep learning models.

Keras

Keras is a neural network library written in Python. It is a high-level API (Ap plication Programming Interface) for creating learning models, providing a consistent syntax and a simple, modular, and extensible interface for building neural networks.

Therefore, TensorFlow and Keras combine perfectly, providing a powerful, easy-to use, and fast execution tandem. Keras can also be used in combination with other frameworks, such as the Microsoft Cognitive Toolkit or Theano (Keras, 2020).

Results

Our research has yielded the following results and contributions:

1. Several datasets for different types of food and their corresponding maturity stages.
2. A food detection model.
3. A maturity degree identification model.
4. In the following subsections, we will describe each contribution in more detail.

Datasets

As previously mentioned, we focused on three types of food that are among the top 5 most wasted in Mexico according to EFE (2017): mangoes, tomatoes, and potatoes. We collected images of these foods at different stages of maturity to create separate datasets for each type of food. In total, we created four datasets: one for food detection and one for each of the three food maturity stages.

The dataset used for the food detection model comprises 1500 labeled images. Most of these images were obtained from the internet, while the rest were taken by us using a cell phone camera. Each image was manually tagged. The source and number of images for each food are summarized in Table 1.

Separate datasets for each maturity stage were utilized to individually train specific maturity identification models. Each model was trained to identify the maturity stage of a specific type of food.

Table 1. Description of the dataset used to build the Food detector model

Source of images	Food	Total	dataset size
Downloaded from the internet	Mango	450	1500
	Tomato	350	
	Potato	300	
	Combined	100	
Own images	Mango	50	
	Tomato	100	
	Potato	100	
	Combined	50	

Note: The row Combined refer to images where appear the 3 foods.

The datasets used to build the models for determining the degree of maturity for each food consisted of at least 800 images, with only four stages of maturity considered for each food type: the maturation stage where the food may not be ready for immediate consumption, the ripening stage where the food is consumable, the senescence stage where the food loses its texture and flavor but is still consumable, and the senescence end stage where the food is no longer consumable and has reached its expiration date. All images in this case were owned by us, and each image was manually tagged. The quantities of the images for each food are summarized in Table 2.

Construction of deep learning models

We utilized the powerful and efficient RetinaNet architecture for accurate object detection across various sizes and classes in images. The architecture comprises a backbone network, a feature pyramid network (FPN), and task-specific subnetworks for regression and classification. The backbone network extracts high-level features from the input image using a pre-trained convolutional neural network (CNN) like ResNet or VGG. It leverages its knowledge from a large dataset of general image recognition tasks. The FPN enhances feature representation by merging multi-scale features from different levels of the backbone network. This creates a feature pyramid with rich semantic information at multiple resolutions, enabling object detection at different scales. The regression subnetwork predicts the bounding box coordinates (x, y, width, height) for each anchor box using the FPN's feature maps. Then the classification subnetwork determines the probability of each anchor box belonging to different object classes. During training, the model employs a combination of loss functions. Focal loss is used for computing the classification loss, addressing the class imbalance issue in object detection. Finally, the regression loss is calculated using the robust smooth L1 loss, accommodating outliers.

In our research, we modified the initial architecture of RetinaNet by only changing the last layer, which determines the number of classes to detect. For the food detector model, we adjusted it to three classes: potato, tomato, and mango. XML files were generated and combined into a single CSV file using a Python script for training. Similarly, for each maturity degree identification model, we used four classes: maturation, ripening, senescence, and senescence end. The training process followed the same approach as the food detector model.

By replacing the classification layer, we can customize the predictions of the model to the specific object categories we want to detect. When working with a pre-trained RetinaNet model trained on a general object detection dataset with diverse classes, we can modify the classification layer to only predict the relevant classes for our task. This enables the model to focus on learning the specific object categories of interest, potentially enhancing performance and reducing inference time.

Modifying the classification layer also helps address class imbalance issues that may arise when the distribution of object classes in our target dataset differs significantly from the original training dataset.

Changing only the classification layer offers a flexible and efficient approach to adapt the RetinaNet architecture for specific object detection tasks. It allows us to leverage the pre-trained backbone network and feature pyramid while tailoring the model's output to meet our specific application requirements.

Table 2. Description of the dataset images used to build the Mango, Tomato, and Potato Maturity Identifier Model

Food	Degree of maturity	Total	Dataset size
Mango	Maturation	50	850
	Ripening	350	
	Senescence	250	
	Senescence end	200	
Tomato	Maturation	20	850
	Ripening	350	
	Senescence	220	
	Senescence end	260	
Potato	Maturation	20	800
	Ripening	300	
	Senescence	215	
	Senescence end	265	

https://doi.org/10.22201/fi.25940732e.2025.26.3.019

The output from the Google Colab environment shown in Figure 2 is the summary of the food detector model architecture. The model architectures for maturity degree identification are similar, except for the number of output layers.

The models are trained using a technique called convolution, where one matrix is treated by another known as a ”kernel”. The convolution matrix filter uses the image as the first matrix to be treated, as shown in Figure 2A. The image is a two-dimensional collection of pixels in rectangular coordinates, and the filter successively examines each pixel of the image to obtain precise information, as seen in Figure 2B. This information is then sent to the classification subnetwork, where it makes a prediction of an object, as shown in Figure 2C. Finally, the obtained results are sent to the regression subnet to verify labels and detected objects, as shown in Figure 2D.

This process is repeated for each image in the training dataset, and once it finishes with all of them, an epoch is completed. The advantage of RetinaNet is that it saves one model per epoch, allowing the best-performing model to be reused for retraining with a larger dataset when required (Figure 3).

Evaluation of deep learning models

In order to efficiently evaluate the models, the 10-fold cross-validation technique was used. A script developed in Python randomly saves 20 % of the images from each dataset in a folder, which belongs to the test dataset for evaluating the model. The remaining 80 % of images were used for the training process. Once the model finishes training, the results of the evaluation with the test images are entered as vectors into a small Python script to obtain the F1-score, Recall, Precision, and Accuracy metrics. This was done 10 times for each model to have an evaluation of their performance.

As mentioned earlier, once the training of a model is complete, the images of the test dataset are passed to the trained model so that the model assigns a class to each image in the test dataset. Finally, the value established by the model is compared with the actual value of the image, and with the help of the sklearn library, the evaluation metrics described are calculated.

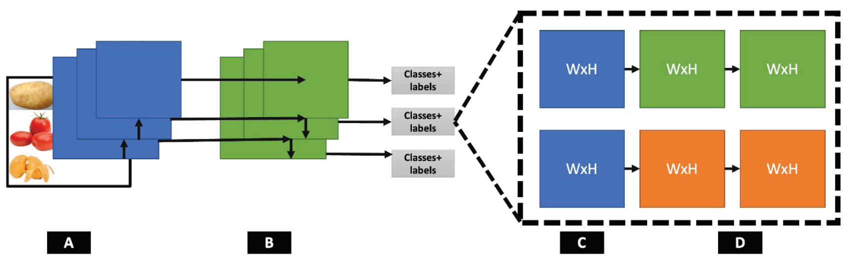

Figure 2. Summary of the Food identification model architecture

```
Root attributes:
  backend: tensorflow
  keras_version: 2.4.0
training_config: {"loss": {"regression": "_smooth_l1", "
                                    classification": "_focal"}, "
                                    metrics": null, "
                                    weighted_metrics": null, "
                                    loss_weights": null, "
                                    optimizer_config": {"class_name
                                    ": "Adam", "config": {"name": "
                                    Adam", "clipnorm": 0.001, "
                                    learning_rate": 9.
                                    99999993922529e-09, "decay": 0.
                                    0, "beta_1": 0.8999999761581421
                                    , "beta_2": 0.9990000128746033,
                                     "epsilon": 1e-07, "amsgrad":
                                    false}}}

classification layer
(None, None, 3)
```

Figure 3. RetinaNet architecture used in our research

Effectiveness of the food detector model

To evaluate the performance of the food detector model, the confusion matrix was calculated as follows:

A real positive if the model identified an object that was actually in the image. A real negative if there was no specific food in the image, and the model did not detect it.

A type II or false negative error if there was a specific food in the image, but the model did not identify it.

A type I or false positive error if there was no specific food in the image, but the model detected it.

The evaluation measures of the effectiveness of the model were then computed, and the results are presented in Table 3.

Table 3. Food detector model's performance

	Precision	Recall	f1 Score		Support
Tomato	0.9200	0.7188	0.8070		32
Mango	0.8000	0.7500	0.7742		48
Potato	0.6410	0.7500	0.7353		29
accuracy				0.7706	109
macro avg	0.7870	0.7769	0.7722		109
weighted avg	0.7929	0.7706	0.7735		109

As shown in Table 3, the Precision value indicates that the model performs well in distinguishing between Tomato and Mango, but is not as effective with Potato. A good Precision value indicates that the model may be labeling False Positives as True Positives, which can increase the value of this measure, but it is not desirable to label examples as True Positives when they are not. The Recall measure gives the effective ness of how well the model identifies True Positives, and is similar in all three classes, indicating an average effectiveness of 74 % in correctly identifying the class of a food (Mango, Tomato, or Potato). The F1 Score remains above 70 %, indicating that both measures are balanced, and thus we have an acceptable effectiveness in distinguishing between the three foods.

Furthermore, the model was trained on more images of Tomato and Mango than Potato, making the weighted avg measure more significant than the macro avg measure. As seen in the results, the weighted measure shows a value higher than 70 %, indicating that despite the class imbalance, the model is effective in distinguishing between the foods.

Effectiveness of the maturity degree identification model

To evaluate the performance of the maturity degree identification model, a confusion matrix was calculated based on the following criteria:

- Whether the model correctly identified a degree of maturity that corresponds to the image, which is considered a true positive.
- If there was no specific degree of maturity in the image and the model did not detect it, which is considered a true negative.
- If there is a specific degree of maturity in the image and the model did not identify it, which is a type II error or false negative.
- If there is no specific degree of maturity in the image, but the model detected it, which is a type I error or false positive.

Table 4 presents the evaluation measures of the effectiveness of each of the three maturity level identification models.

The results presented in Table 4 provide insights into the model's performance in identifying the maturity levels of tomatoes, potatoes, and mangoes. Despite the small number of test images for tomatoes, the model demonstrated overall effectiveness in detecting maturity levels, as indicated by the accuracy score. The recall value further supported this, showing correct identification of maturity levels above 75 %, with exceptions in specific classes. The F1 Score maintained a balanced measure, reflecting acceptable effectiveness in identifying maturity levels. The weighted avg measurement accounted for the class imbalance and determined that the model could identify tomato maturity levels with an average above 70 %. Although the macro avg measure was low, its impact was limited due to the class imbalance.

Regarding potato maturity levels, the model exhibited high precision but suffered from incorrectly classifying false positives as true positives. This issue was evident in the precision measure for maturation and senescence end classes, which can lead to mis leading results. However, the recall values indicated that the model generally performed well in identifying potato maturity levels, with an average value above 80 %. The F1 Score provided a balanced measure, indicating the model's effectiveness in identifying three out of the four degrees of maturity. Once again, the class imbalance was apparent, emphasizing the importance of the weighted avg measure, which demonstrated the model's performance in identifying potato maturity levels with a measure of over 80 %. The macro avg mea-

https://doi.org/10.22201/fi.25940732e.2025.26.3.019

Table 4. Tomato Maturity Level Identification Models' performance

		Precision	Recall	f1 Score		Support
	Maturation	0.000	0.000	0.000		4
	Ripening	0.7655	0.9142	0.8332		35
	Senescence	0.7678	0.9749	0.8558		24
Tomato	Senescence end	0.5512	0.2066	0.3810		15
	accuracy				0.7499	78
	macro avg	0.5553	0.5452	0.5239		78
	weighted avg	0.7166	0.7692	0.7197		78
	Maturation	1.000	0.2500	0.4000		4
	Ripening	0.7667	0.7783	0.7964		30
	Senescence	0.7647	0.9286	0.8387		28
Potato	Senescence end	1.000	0.8636	0.9268		22
	accuracy				0.8214	84
	macro avg	0.8828	0.7022	0.7331		84
	weighted avg	0.8382	0.8214	0.8152		84
	Maturation	0.9200	0.3750	0.6667		8
	Ripening	0.9091	0.8889	0.8989		45
	Senescence	0.6207	0.75	0.6792		24
Mango	Senescence end	0.6774	0.5676	0.6176		37
	accuracy				0.7176	114
	macro avg	0.7700	0.6502	0.6695		114
	weighted avg	0.7390	0.7193	0.7145		114

sure also confirmed the effectiveness of the model, despite the class imbalance.

For mango maturity levels, the precision measure suggested the model's proficiency, particularly in two classes. However, the issue of false positives being considered as true positives persisted, resulting in inflated values. This issue was reflected in the re call measure, notably in the senescence end class, where the measurement considerably decreased. The average recall value of 70 % indicated that the model could identify 70 % of the detected maturity degrees. The F1 Score maintained an average of 70 %, signifying high effectiveness in identifying the ripening class but lower performance in other classes. Nonetheless, the average F1 Score aligned with the expected effectiveness of the mentioned class. The macro avg measure demonstrated similar values, highlighting the class imbalance, while the weighted avg measure provided a more appropriate parameter, indicating an average of 71 % and confirming the model's correct identification of maturity degrees despite the varying number of test images in each class.

In summary, the models displayed varying performance in identifying the maturity levels of tomatoes, potatoes, and mangoes, with some challenges related to false positives. However, they generally achieved satisfactory results, particularly in the majority classes, emphasizing their effectiveness in identifying maturity levels. The class imbalance was evident in all models, underscoring the significance of the weighted avg measure in evaluating their performance.

Impacts and limitations

The implementation of deep learning models to identify food maturity stages can yield significant impacts on supermarkets and farmers. For supermarkets, accurately identifying food maturity stages in real-time can optimize inventory management, leading to reduced food spoilage and waste. Supermarkets can benefit financially by lowering losses associated with unsold perishable products, thus increasing profitability and operational efficiency. Additionally, improved maturity identification contributes positively to sustainability goals and enhances corporate social responsibility profiles. For farmers, using these models can improve decision-making during harvesting and distribution processes. Identifying optimal maturity points can help farmers ensure that products reach the market at ideal freshness, which can command better prices and reduce rejections from buyers due to suboptimal maturity. Finally, the technology supports farmers in decreasing losses from prematurely harvested or overly matured crops, enhancing their economic resilience and market competitiveness.

It is important to acknowledge that the variability in lighting conditions presented a significant limitation during data collection, potentially affecting deep learning model performance due to their reliance on visual features for food maturity classification. Inconsistent illumination, including shadows, overexposure, or inadequate lighting, can lead to misclassification. While this is a common challenge, our results are validated by extensive preprocessing, including image augmentations like rotations, scaling, and lighting adjustments. Nonetheless, acknowledging this noise source is crucial for developing and deploying deep learning applications in real-world scenarios. Future implementations should prioritize robust preprocessing and normalization to mitigate its impact.

Conclusion and future work

We have demonstrated the feasibility of using deep learning models to determine the maturity levels of three types of food that are frequently wasted in Mexico. This approach is novel in research areas dedicated to developing methods, techniques, or technologies for avoiding or reducing food waste.

The food identification model successfully identified the selected foods (potato, mango, and tomato) with an effectiveness rate of 77 %. The maturity level identification models achieved an effectiveness rate of 70 %, which is similar to results reported in other studies. It is worth mentioning that these outcomes were largely attributed to the size of the dataset and the information obtained from training images. However, it should be noted that performance can be improved by increasing the dataset size.

Considering the obtained results and the discussion surrounding the RetinaNet architecture, a future extension of the research could involve identifying not only the three selected foods but also other commonly wasted food items. This extension could significantly contribute to waste reduction on a global scale.

In terms of future work, the authors propose the development of a more comprehensive system to have a significant impact on reducing food waste. This system would leverage technologies like the Internet of Things (IoT), cloud computing, and mobile applications to automate the process. By incorporating a real-time camera and utilizing cloud hosting, the system could accurately identify the moments of transition between different maturity stages of food. Real-time notifications and suggestions regarding food maturation could be provided to users with internet access.

To improve the performance and accuracy of the models, a larger and more diverse dataset is recommended. Including additional images with various textures, shapes, and colors would provide a comprehensive understanding of the different maturity stages of food, leading to reduced error margins and increased efficiency. Furthermore, exploring the use of alternative pre-trained models with different architectures or developing a custom model could be beneficial in enhancing the accuracy and effectiveness of the system.

References

Balas, V. E. (2019). *Handbook of deep learning applications*. Springer International Publishing.

Barrios, J. (2019). *La matriz de confusión y sus métricas*. In Health BIG DATA, Ciencia de datos, Informática Médica, Inteligencia Artificial, Machinelearning.

Berruguete, P. G. (2019). Descomposición de alimentos. *Seguridad Alimentaria*, 49, 69-78.

Bhargava, A., & Bansal, A. (2021). Fruits and vegetables quality evaluation using computer vision: A review. *Journal of King Saud University-Computer and Information Sciences*, 33(3), 243-257. https://doi.org/10.1016/j.jksuci.2018.06.002

Carvajal-Millan, E., Yaser, A. Z., & Haghi, A. K. (2021). *In natural food products and waste recovery: Healthy foods, nutrition design, and extraction of valuable compounds*. Apple Academic Press.

Chai, J., Zeng, H., Li, A., & Ngai, E. W. T. (2021). Deep learning in computer vision: A critical review of emerging techniques and application scenarios. *Machine Learning with Applications*, 6, 100134. https://doi.org/10.1016/j.mlwa.2021.100134

Chasiotis, V., Tzempelikos, D., Filios, A., & Moustris, K. P. (2020). Artificial neural network modelling of moisture content evolution for convective drying of cylindrical quince slices. *Computers and Electronics in Agriculture*, 172, 105074. http://dx.doi.org/10.1016/j.compag.2019.105074

Dang, N. T., Vo, M. T., Nguyen, T. D., & Dao, S. V. T. (2019). Analysis on mangoes weight estimation problem using neural network. 19th International Symposium on Communications and Information Technologies (ISCIT), 559-562. Retrieved on https://doi.org/10.1109/ISCIT.2019.8905118

EFE, A. (2017). Banco Mundial: México despilfarra el 34 % de su producción alimentaria. Retrieved on https://www.efe.com/efe/america/mexico/banco-mundial-mexico-despilfarra-el-34-de-su-produccion-alimentaria/50000545-3461427

Farinella, L., Fernandes, E., Michener, N., Polimeni, M., & Vesonder, G. (2020). Identification of food waste through object recognition. 11th IEEE Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON). Retrieved on https://doi.org/10.1109/uemcon51285.2020.9298165

Hathi, H. S., Parmar, D. L., Bedva, S. M., Purohit, S. A., & Patel, D. B. (2020). Maturity indices of fruits and vegetables. *Vigyan Varta*, 1(8), 38-42.

Kakani, V., Nguyen, V. H., Kumar, B. P., Kim, H., & Pasupuleti, V. R. (2020). A critical review on computer vision and artificial intelligence in food industry. *Journal of Agriculture and Food Research*, 2, 100033. https://doi.org/10.1016/j.jafr.2020.100033

Keras. (2020). Keras, the Python deep learning API.

Koszela, K., Łukomski, M., Mueller, W., Górna, K., Okoń, P., Boniecki, P., Zaborowicz, M., & Wojcieszak, D. (2017). Classification of dried vegetables using computer image analysis and artificial neural networks. In Falco C. M. & Jiang X. (Eds.). Ninth International Conference on Digital Image Processing (ICDIP 2017), Vol. 10420, 650-656. SPIE. Retrieved on https://doi.org/10.1117/12.2281718

Kumar, I., Rawat, J., Mohd, N., & Husain, S. (2021). Opportunities of Artificial Intelligence and machine learning in the Food Industry. *Journal of Food Quality*, 2021, 1-10. https://doi.org/10.1155/2021/4535567

Kumar, T. B., Prashar, D., Vaidya, G., Kumar, V., Kumar, S. D., & Sammy, F. (2022). A novel model to detect and classify fresh and damaged fruits to reduce food waste using a deep learning technique. *Journal of Food Quality*, 2022, 1-8. https://doi.org/10.1155/2022/4661108

Kunszabó, A., Szakos, D., & Kasza, G. (2019). Food waste-A general overview and possible solutions. *Hungarian Agricultural Research*, 28(3), 14-19.

Li, Y. (2022). Research and application of deep learning in image recognition. IEEE 2nd International Conference on Power, Electronics and Computer Applications (ICPECA), 994-999. Retrieved on https://doi.org/10.1109/ICPECA53709.2022.9718847

Lin, T., Goyal, P., Girshick, R. B., He, K., & Dollár, P. (2017). Focal loss for dense object detection. CoRR, abs/1708.02002. Retrieved on http://arxiv.org/abs/1708.02002

Liu, J., Liu, L., Guo, W., Fu, M., Yang, M., Huang, S., Zhang, F., & Liu, Y. (2019). A new methodology for sensory quality assessment of garlic based on metabolomics and an artificial neural network. *RSC Advances*, 9(31), 17754-17765. https://doi.org/10.1039/C9RA01978B

Moreno, C. (2016). *Valoremos los alimentos, evitemos pérdidas y desperdicios*. Biblioteca Em Saúde.

Morone, P., Koutinas, A., Gathergood, N., Arshadi, M., & Matharu, A. (2019). Food waste: Challenges and opportunities for enhancing the emerging bio-economy. *Journal of Cleaner Production*, 221, 10-16. https://doi.org/10.1016/j.jclepro.2019.02.258

Mozos, E. A. de los, Badurdeen, F., & Dossou, P.-E. (2020). Sustainable consumption by reducing food waste: A review of the current state and directions for future research. *Procedia Manufacturing,* 51, 1791-1798. https://doi.org/10.1016/j.promfg.2020.10.249

Panda, S., & Dwivedi, M. (2020). Minimizing food wastage using machine learning: A novel approach, 465-473. https://doi.org/10.1007/978-981-13-9282-5_44

Patel, P., & Thakkar, A. (2020). The upsurge of deep learning for computer vision applications. *International Journal of Electrical and Computer Engineering (IJECE),* 10, 538. https://doi.org/10.11591/ijece.v10i1.pp538-548

Reynolds, C. (2020). *In Routledge Handbook of Food Waste*. 1a ed. Routledge.

Tensorflow. (2019). ¿Por qué Tensorflow?

Umeki, H. (2014). What is Deep Learning?

UN. (2016). The zero hunger challenge-Advisory notes for action.

Zhang, S., Chen, Y., Yang, Z., & Gong, H. (2021). Computer vision based two-stage waste recognition-retrieval algorithm for waste classification. *Resources, Conservation and Recycling*, 169, 105543. https://doi.org/10.1016/j.resconrec.2021.105543

How to cite:

Cuaya-Simbro, G., & Cuevas-Ordoñez, D. (2025). Deep learning to identifying food maturity for prevent food waste. *Ingeniería Investigación y Tecnología*, 26(03), 1-12. https://doi.org/10.22201/fi.25940732e.2025.26.3.019