

## Clasificación para frutos de fresa con CNN

César Augusto Pilon Alcalá, Juan Carlos Olguín Rojas

Universidad Autónoma Chapingo,  
Departamento de Ingeniería Mecánica Agrícola,  
México

jolguinr@chapingo.mx

**Resumen.** El uso de la inteligencia artificial (IA), en particular de las redes neuronales convolucionales (RNN), ha revolucionado la capacidad de clasificar objetos y productos con precisión, incluso para personas sin experiencia en el campo. En el caso específico de la clasificación de fresas *Fragaria* × *ananassa* en su estado óptimo, las RNN pueden analizar imágenes y discernir características que indican su frescura, madurez y calidad. Este estudio busca encontrar una arquitectura óptima para la clasificación de fresas basada en imágenes. Para lograr esto, se analizaron tres arquitecturas de redes neuronales convolucionales (LeNet5, VGG16 y VGG19). Se llevaron a cabo varios procedimientos, que consistieron en combinaciones de redes con sus respectivos hiperparámetros. Se utilizaron diferentes tasas de aprendizaje (en un rango de  $1e-8$  a  $0.1$ ), cuatro optimizadores distintos (Adagrad, Adam, RMSProp y SGD), y se emplearon técnicas de entrenamiento (extracción de características) para la clasificación del objeto de estudio (*Fragaria* × *ananassa*). Cada uno de estos procedimientos fue evaluado para determinar y comparar su rendimiento. La arquitectura que mejor desempeño obtuvo fue LeNet5, con el optimizador RMSProp y una tasa de aprendizaje de  $0.001$ , logrando una precisión del  $90.6\%$ .

**Palabras clave:** Redes neuronales convolucionales, *fragaria* × *ananassa*, LeNet5, VGG16, VGG19.

## Classification of Strawberry Fruits Using CNN

**Abstract.** The use of artificial intelligence (AI), particularly convolutional neural networks (CNNs), has revolutionized the ability to accurately classify objects and products, even for individuals without expertise in the field. In the specific case of classifying strawberries *Fragaria* × *ananassa* at their optimal state, CNNs can analyze images and discern features indicating their freshness, ripeness, and quality. This study seeks to find an optimal architecture for strawberry classification based on images. To achieve this, three convolutional neural network architectures (LeNet5, VGG16, and VGG19) were analyzed. Various procedures were carried out, consisting of combinations of networks with their respective hyperparameters. Different learning rates were used (ranging from  $1e-8$  to  $0.1$ ), four distinct optimizers (Adagrad, Adam, RMSProp, and SGD), and

training techniques (feature extraction) were employed for the classification of the study subject (*Fragaria × ana-nassa*). Each of these procedures was evaluated to determine and compare their performance. The architecture that performed best was LeNet5, with the RMSProp optimizer and a learning rate of 0.001, achieving an accuracy of 90.6%.

**Keywords:** Convolutional neural networks, *fragaria × ananassa*, LeNet5, VGG16, VGG19.

## 1. Introducción

La fresa es una fruta de gran importancia y rentabilidad. Con su alto contenido de vitamina C y valor nutricional, es muy versátil en su uso, ya sea para el consumo directo, la decoración de postres o la elaboración de jugos, entre otros usos.

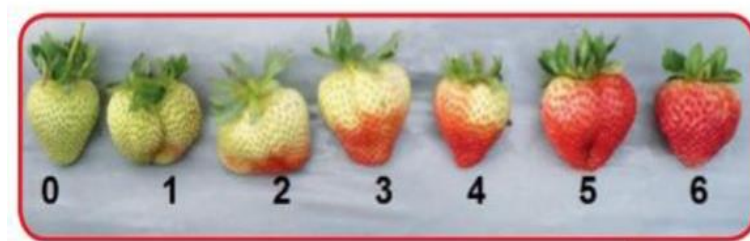
Sin embargo, su corta vida útil después de la cosecha y su susceptibilidad a lesiones cuando están muy maduras representan desafíos significativos para su almacenamiento y venta [18]. En el 2022 en México se ocupó el octavo lugar en producción de fresa con 568271.93 toneladas [7], y los estados con mayor producción son Michoacán con 354,047.99 t, Guanajuato con 96,119.00 t y Baja California con 93,179.00 t. [13].

Las fresas pertenecen a la familia de las rosáceas y pasan por el ciclo de floración, fructificación y maduración. No hay criterios oficiales establecidos para la clasificación de las diferentes etapas de madurez. El método tradicional para determinar las etapas de madurez de las fresas implica evaluar manualmente su apariencia, color, textura, sabor y firmeza. En consecuencia, resulta fundamental establecer un estándar para la clasificación de las etapas de madurez y desarrollar un método económico y eficaz para identificar con rapidez y precisión cada una de estas [20].

La duración de la etapa de consumo de las fresas después de ser cosechadas es breve, y demasiado maduras, tienden a dañarse, por eso es complicado almacenarlas y comercializarlas. Por lo tanto, es especialmente crucial seleccionar el momento adecuado para recolectarlas, considerando el periodo específico de almacenamiento. Porque, para que la fresa casi madura pase a ser un fruto maduro requiere aproximadamente de 5 a 10 días [18].

La selección de los frutos de fresa en el estado óptimo de maduración garantiza un producto de excelentes condiciones y buena vida útil. Si se cosechan frutos fisiológicamente inmaduros no se desarrolla su color y sabor, convirtiéndose en un producto de calidad inferior. Si se cosecha un fruto maduro esté tendrá una vida comercial corta y se fermentará fácilmente dañando los frutos a su alrededor [4]. Una alternativa viable a este problema es la implementación de técnicas de visión computacional que no son invasivas.

Estudios previos como el de Yuanyuan Shao y su equipo en [12] utilizaron un sistema portátil de imágenes hiperespectrales para la adquisición de espectros de fresas en el campo y en interiores en 3 etapas de madurez: maduras, semimaduras y verdes. Por otro lado, Xiao-Qin Yue y su grupo de investigación [18] lograron reconocer mediante teléfonos inteligentes (smartphones) y con imágenes de fresas a longitudes de



**Fig. 1.** Tabla de color de la fresa [4].

onda de 535 nm y 670 nm 3 etapas: madura, casi madura e inmadura y Xiaofen Du y sus colaboradores [6] clasificaron mediante una nariz electrónica (e-nose) compuesta por 18 sensores de gases de óxido metálico diferentes, para caracterizar los patrones volátiles de las fresas en 5 etapas de desarrollo: blanco, medio rojo, tres cuartos de rojo, completamente maduro y sobremaduro.

El procesamiento de imágenes utilizando métodos de aprendizaje profundo está siendo investigado para aplicaciones agrícolas debido al aumento de la velocidad de procesamiento y los avances en algoritmos informáticos [19]. Se han utilizado diversas arquitecturas de redes neuronales convolucionales, como YOLOv3 para clasificación y detección de las etapas de maduración de fresa [20], Mask R-CNN para detectar una máscara pixel a pixel de naranja [8] y magulladuras en imágenes de fresas capturadas por una cámara a color bajo luz incandescente y ultravioleta (UV) [19] y YOLOv5 para el reconocimiento en tiempo real del tallo/cáliz de las manzanas [16].

Una CNN (CNN del inglés Convolutional Neural Networks) es una red neuronal con varios tipos de capas especiales. Hoy en día este tipo de red está siendo muy usada por la industria para diversas tareas, especialmente de visión por computacional [15].

La falta de información relacionada con la clasificación de fresas con redes neuronales ha motivado la creación del presente trabajo, por lo que el objetivo de este escrito es encontrar una arquitectura de red neuronal convolucional para la clasificación de fresa en tres etapas de maduración (inmadura, madura y muy madura), para lograrlo, se considerarán algunas arquitecturas CNN que han reportado buenos resultados en diversas aplicaciones.

## 2. Materiales y métodos

### 2.1. Arquitecturas CNN utilizadas

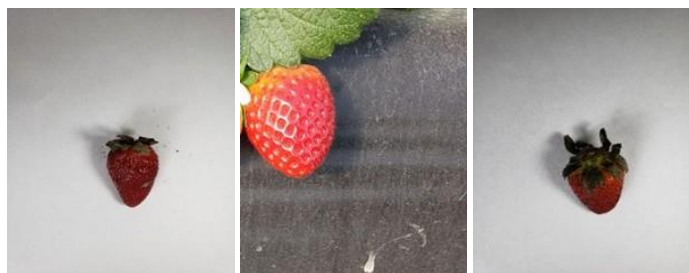
Para lograr una clasificación de fresa en 3 etapas de madurez (inmadura, madura y muy madura) de manera no invasiva, se consideraron 3 arquitecturas; LeNet5, basada en la propuesta de Yann LeCun y colaboradores [11]; VGG16 y VGG19 de Karen Simonyan y Andrew Zisserman [14]. Ampliamente reconocidas y utilizadas en diversas áreas, como la robótica, seguridad, agricultura [17] y medicina [10] y en comparación con arquitecturas como Resnet 50, Resnet 101 o Inception v4, son arquitecturas que



**Fig. 2.** Ejemplo de fresa inmadura.



**Fig. 3.** Ejemplo de fresa madura.

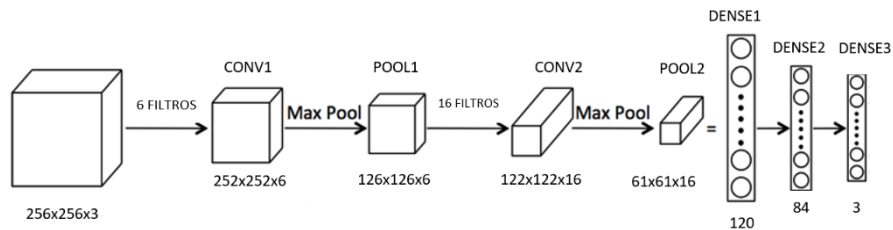


**Fig. 4.** Ejemplo de fresa muy madura.

emplean un menor número total de parámetros, lo que permite obtener un rendimiento aceptable en implementaciones de baja eficiencia computacional.

## 2.2. Conjunto de datos

Una de las actividades para realizar esta investigación, fue la creación de un conjunto de imágenes de fresa para identificar las 3 categorías a clasificar (inmadura, madura y muy madura). Para lograr esto, se realizó una búsqueda de imágenes de fresa en la web de diferentes tamaños de imagen y en 3 canales (RGB), las imágenes encontradas se recortaron a conveniencia y se categorizaron según la norma NMX-FF-062-SCFI-2002 [5] que establece que las fresas maduras deben tener una coloración roja que se extiende desde el ápice hasta la base del pedúnculo y cubrir al menos el 50% de la superficie y



**Fig. 5.** Esquema ilustrativo de la arquitectura LeNet5.

el Instituto Tecnológico Superior de Coacolmán [4] (ver Fig. 1). Las fresas con características similares a 0, 1 y 2 se categorizaron como inmaduras; 3, 4 y 5 como maduras y 6 como muy maduras.

Se obtuvo un total de 1524 imágenes con diferentes dimensiones. Se generó un conjunto de datos de 710 imágenes para fresas en estado inmaduro, 453 para fresas maduras y 361 para fresas muy maduras. El 70% de cada conjunto de imágenes mencionado se empleó para el entrenamiento, el 15% para la validación y el 15% para las pruebas.

En las Fig. 2, 3 y 4 se muestra un ejemplo de las imágenes utilizadas en este estudio.

### 2.3. Arquitectura LeNet5

La arquitectura LeNet5 fue creada por Yann LeCun en 1998 y es ampliamente utilizada para el reconocimiento de dígitos escritos a mano (MNIST) [9], está compuesta por 2 capas convolucionales, 2 capas de agrupamiento máximo (max pooling) y 3 capas densas (dense). Para este trabajo se usó esta arquitectura y se propusieron 2 capas convolucionales de 6 y 16 filtros respectivamente con un tamaño de 5x5; 2 capas de agrupamiento máximo de 2x2 y 3 capas densas; en el clasificador se tiene una capa de 120 neuronas con función de activación relu; otra de 84 neuronas con función de activación relu y la última es de 3 neuronas con función de activación softmax (ver Fig. 5).

### 2.4. Arquitectura VGGNet

También, se utilizó la arquitectura VGG16 (ver Fig. 6) y VGG19 que originalmente fueron preentrenadas con Imagenet. Para cada una de estas, se empleó aprendizaje por transferencia, se usó la técnica de extracción de características (feature extraction) descongelando los pesos del bloque convolucional 5 de cada arquitectura y se reentrenaron 1 y 2 capas convolucionales, finalmente en el clasificador se tienen 2 capas densas de 512 neuronas con función de activación relu y una capa final de 3 neuronas con función de activación softmax.

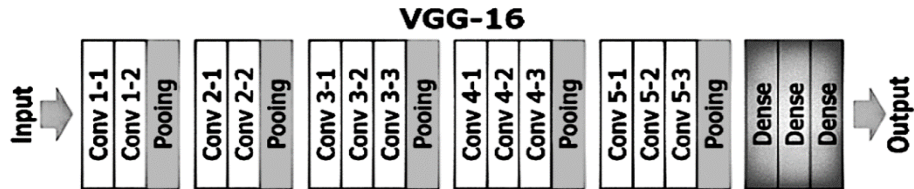


Fig. 6. Esquema ilustrativo de la arquitectura VGG16.

## 2.5. Diseño del experimento

Para los ensayos, se consideró la búsqueda por rejilla y se propusieron una combinación de hiperparámetros para medir el mejor rendimiento en clasificación, al final de cada entrenamiento, se seleccionaron los mejores rendimientos. Los factores involucrados en cada tratamiento fueron la arquitectura, el optimizador, la tasa de aprendizaje, y en el caso de las arquitecturas VGG, además, el número de capas finales a reentrenar.

Debido al amplio número de combinaciones en la búsqueda por rejilla, se optó por hacer un diseño de experimento en 2 etapas; en la primera se entrenaron las arquitecturas solo a 30 épocas, esto con el fin de obtener las mejores tasas de aprendizaje y reducir el espacio de búsqueda, posteriormente en la segunda etapa, se realizaron nuevamente los experimentos con las mejores tasas de aprendizaje, encontradas anteriormente, pero a 100 épocas. A continuación, se describirá con detalle cada una de las etapas.

## 2.6. Etapa 1. Búsqueda rápida de tasa de aprendizaje

El objetivo de esta búsqueda fue encontrar un intervalo de valores para la tasa de aprendizaje de los mejores rendimientos. Para ello se hicieron ensayos con valores diferentes de tasa de aprendizaje entre; 0.1 y 1e-8, se utilizaron 4 optimizadores diferentes; Adam, Adagrad, RMSProp y SGD para las arquitecturas Lenet5, VGG16 y VGG19; para las 2 últimas, se utilizó la técnica de extracción de características, la cual consistió en descongelar solo las capas convolucionales 1 y 2 y del bloque 5 (block5\_conv1 y block5\_conv2), y en otro ensayo se descongeló solo la capa 3 del bloque 5 (block\_conv3) para cada arquitectura VGGNet. En esta etapa las mejores tasas de aprendizaje fueron 0.001, 0.0001, 0.00001, 0.000001 y 0.0000001.

## 2.7. Etapa 2. Búsqueda complete

El objetivo de esta fase fue encontrar la arquitectura, optimizador y tasa de aprendizaje que mejor se desempeñe para este problema específico. Las arquitecturas que se utilizaron fueron LeNet5, VGG16 y VGG19, los optimizadores fueron los 4 mencionados en la etapa 1, se usaron las 5 tasas de aprendizaje obtenidas en la etapa

**Tabla 1.** 10 mejores rendimientos de 100 tratamientos realizados de la etapa 2. Para obtener la exactitud se promedió la exactitud de entrenamiento, validación y prueba de cada tratamiento. Los tratamientos fueron realizados con 100 épocas.

| Tratamiento | Arquitectura                 | Optimizador | Tasa de aprendizaje | Pesos iniciales | Exactitud |
|-------------|------------------------------|-------------|---------------------|-----------------|-----------|
| 6           | LeNet 5                      | Adam        | 0.001               | Aleatorio       | 0.906     |
| 7           | LeNet 5                      | Adam        | 0.0001              | Aleatorio       | 0.908     |
| 11          | LeNet 5                      | RMSProp     | 0.001               | Aleatorio       | 0.906     |
| 16          | LeNet 5                      | SGD         | 0.001               | Aleatorio       | 0.907     |
| 26          | VGG16<br>reentrenando 1 capa | Adam        | 0.001               | ImageNet        | 0.904     |
| 27          | VGG16<br>reentrenando 1 capa | Adam        | 0.0001              | ImageNet        | 0.904     |
| 28          | VGG16<br>reentrenando 1 capa | Adam        | 0.00001             | ImageNet        | 0.909     |
| 46          | VGG16<br>reentrenando 2 capa | Adam        | 0.001               | ImageNet        | 0.906     |
| 48          | VGG16<br>reentrenando 2 capa | Adam        | 0.00001             | ImageNet        | 0.897     |
| 51          | VGG16<br>reentrenando 2 capa | RMSProp     | 0.001               | ImageNet        | 0.895     |

anterior y las técnicas de entrenamiento fueron 2 desde cero para la arquitectura LeNet5; mientras que, para VGG16 y VGG19 se optó por usar aprendizaje por transferencia con la técnica de extracción de características; un ensayo fue descongelando solo la capa convolucional 3 del bloque 5 (block\_conv3), como se explicó en la etapa 1 y otro ensayo donde se descongelaron las capas 1 y 2 del bloque 5 (block5\_conv1 y block5\_conv2) para cada arquitectura VGGNet, haciendo un total del 100 ensayos.

Además, se realizó un reescalado de imágenes a (256, 256, 3) y se estableció un lote por época de 11 imágenes, para el entrenamiento se utilizó Keras, con Python 3.7 y una versión de Tensorflow de 2.1.0

## 2.8. Métricas de evaluación

Para valorar el rendimiento de cada tratamiento se emplearon medidas de precisión, exactitud, puntuación F1 y también se analizó la matriz de confusión, consultar [9].

## 3. Resultados

Los mejores 10 resultados experimentales de la etapa 2, se encuentran en la Tabla 1. Se puede observar que la exactitud se encuentra entre los intervalos de 0.895 a 0.908. y que en los tratamientos 7 y 28, con la arquitectura LeNet5 y VGG16 son los que tienen la mejor exactitud con 0.908 y 0.909 respectivamente.

**Tabla 2.** Comparación de métricas de desempeño de tratamientos 6, 7, 11, 16, 26, 27, 28 y 46. Se utilizaron 54 imágenes para 'Validación fresa muy madura', 68 para 'Validación fresa madura' y 106 para 'Validación fresa inmadura'.

| LeNet5, tratamiento 6                       |           |              |               |
|---------------------------------------------|-----------|--------------|---------------|
| Categoría                                   | Precisión | Sensibilidad | Puntuación F1 |
| Validación fresa muy madura                 | 0.95      | 0.7          | 0.81          |
| Validación fresa madura                     | 0.79      | 0.93         | 0.85          |
| Validación fresa inmadura                   | 0.97      | 0.99         | 0.98          |
| Promedios                                   | 0.903     | 0.873        | 0.880         |
| LeNet5, tratamiento 7                       |           |              |               |
| Categoría                                   | Precisión | Sensibilidad | Puntuación F1 |
| Validación fresa muy madura                 | 0.8       | 0.91         | 0.81          |
| Validación fresa madura                     | 0.86      | 0.81         | 0.83          |
| Validación fresa inmadura                   | 0.99      | 0.96         | 0.98          |
| Promedios                                   | 0.883     | 0.893        | 0.873         |
| LeNet5, tratamiento 11                      |           |              |               |
| Categoría                                   | Precisión | Sensibilidad | Puntuación F1 |
| Validación fresa muy madura                 | 0.94      | 0.81         | 0.87          |
| Validación fresa madura                     | 0.86      | 0.91         | 0.89          |
| Validación fresa inmadura                   | 0.97      | 1            | 0.99          |
| Promedios                                   | 0.923     | 0.907        | 0.917         |
| LeNet5, tratamiento 16                      |           |              |               |
| Categoría                                   | Precisión | Sensibilidad | Puntuación F1 |
| Validación fresa muy madura                 | 0.81      | 0.87         | 0.84          |
| Validación fresa madura                     | 0.89      | 0.82         | 0.85          |
| Validación fresa inmadura                   | 0.99      | 1            | 1             |
| Promedios                                   | 0.897     | 0.897        | 0.897         |
| VGG16 descongelando 1 capa, tratamiento 26  |           |              |               |
| Categoría                                   | Precisión | Sensibilidad | Puntuación F1 |
| Validación fresa muy madura                 | 0.84      | 0.78         | 0.81          |
| Validación fresa madura                     | 0.76      | 0.88         | 0.82          |
| Validación fresa inmadura                   | 0.98      | 0.92         | 0.95          |
| Promedios                                   | 0.860     | 0.860        | 0.860         |
| VGG16 descongelando 1 capa, tratamiento 27  |           |              |               |
| Categoría                                   | Precisión | Sensibilidad | Puntuación F1 |
| Validación fresa muy madura                 | 0.82      | 0.83         | 0.83          |
| Validación fresa madura                     | 0.78      | 0.88         | 0.83          |
| Validación fresa inmadura                   | 0.99      | 0.9          | 0.94          |
| Promedios                                   | 0.863     | 0.870        | 0.867         |
| VGG16 descongelando 1 capa, tratamiento 28  |           |              |               |
| Categoría                                   | Precisión | Sensibilidad | Puntuación F1 |
| Validación fresa muy madura                 | 0.86      | 0.8          | 0.83          |
| Validación fresa madura                     | 0.79      | 0.85         | 0.82          |
| Validación fresa inmadura                   | 0.95      | 0.94         | 0.95          |
| Promedios                                   | 0.867     | 0.863        | 0.867         |
| VGG16 descongelando 2 capas, tratamiento 46 |           |              |               |
| Categoría                                   | Precisión | Sensibilidad | Puntuación F1 |
| Validación fresa muy madura                 | 0.8       | 0.81         | 0.81          |
| Validación fresa madura                     | 0.78      | 0.82         | 0.8           |
| Validación fresa inmadura                   | 0.96      | 0.92         | 0.94          |
| Promedios                                   | 0.847     | 0.850        | 0.850         |

En la Tabla 2 se reportan las comparaciones de las métricas de precisión, sensibilidad y F1 de las arquitecturas con mejor exactitud de la Tabla 1 (tratamientos 6, 7, 11, 16, 26, 27, 28 y 46).

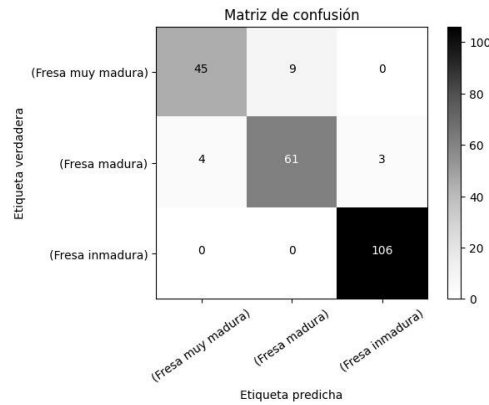
Se determinó que el tratamiento 11 de la Tabla 2 con arquitectura LeNet5, con una tasa de aprendizaje de 0.001 y optimizador RMSProp fue el tratamiento con mejor

desempeño, con una exactitud de 0.906 y los mejores promedios en las métricas de desempeño; 0.923 en precisión, 0.907 en sensibilidad y 0.917 en puntuación F1.

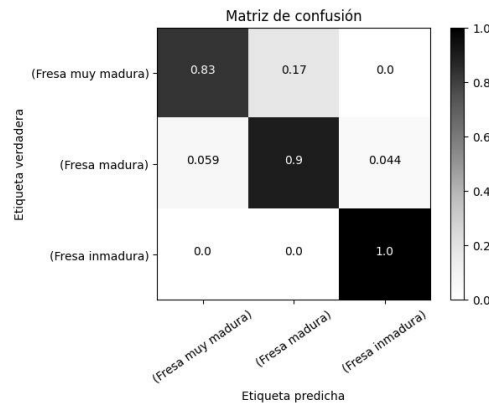
En la evaluación de la matriz de confusión, se utilizaron 228 imágenes y el desempeño en validación se muestra en la Fig. 7 sin normalizar y en la Fig 8 normalizada.

#### 4. Discusión

La arquitectura LeNet5 con una tasa de aprendizaje de 0.001 y un optimizador RMSProp obtuvo un 83% en la fresa muy madura, un 90% en la fresa madura y un 100% en la fresa inmadura en la matriz de confusión y una exactitud de 90.6%. Xiao Qin Yue y colaboradores [18], en su artículo, informaron que con teléfonos inteligentes (smartphones) tiene un 97% en fresa madura, un 87% en fresa casi madura y un 100% en fresa inmadura en su matriz de confusión de testeo, una exactitud de 94.44%, una precisión promedio de 94.85%, una sensibilidad promedio de 94.45% y una puntuación F1 promedio de 94.43%. Yuanyuan Shao y su equipo [12] reportaron que se obtuvo una exactitud en campo de hasta 96.7% usando imágenes hiperespectrales de fresa en 3 etapas de madures (madura, semi madura y no madura) y métodos como LS-SVM (Máquinas de Soporte Vectorial de Mínimos Cuadrados), concluyendo que la imagen hiperespectral puede utilizarse para la evaluación en tiempo real de la madurez de las fresas en el campo. Según Ji-Young Choi y su grupo de investigación [3], desarrollaron modelos de redes neuronales convolucionales (CNNs) para clasificar la calidad externa de fresas (frescas, magulladas o mohosas) utilizando 750 imágenes RGB. Se utilizaron 8 configuraciones para comparar las precisiones de validación según variables como la distribución de datos, el número de imágenes y las épocas de entrenamiento. Se observó un alto rendimiento de aprendizaje en poco tiempo cuando se utilizó el 90% de los datos de imagen para el entrenamiento y la validación. Las métricas de rendimiento fueron exactitud (AC) promedio de 97%, precisión (PR) promedio de 95.6%, especificidad (SP) promedio de 95.6%, sensibilidad (SE) promedio de 97.8% y puntuación F1 promedio de 95.6% del modelo 90% de entrenamiento y 10% de testeo. Además, las CNNs identificaron áreas dañadas en las fresas, mostrando su potencial para el monitoreo no destructivo en la industria alimentaria. Mientras que Liane Angelo Acero y colaboradores [1] emplearon imágenes clasificadas como deseables e indeseables, utilizando 350 imágenes de cada conjunto para el entrenamiento, 200 imágenes para la validación y 100 imágenes para la prueba. El modelo de CNN generado se simuló aplicando épocas (epoch) = 15 y un tamaño de lote (batch size) = 8. Obtuvieron una precisión de entrenamiento del 98.41%, una precisión de validación del 92.75%, y una precisión de prueba del 100%. Por ultimo Hossein Azizi y su equipo [2] utilizaron un conjunto de datos que contiene 800 imágenes de fresas en 4 clases (inmadura, semi-madura, madura y dañada), con 3 escenarios; el primero donde utilizan los datos originales; el segundo empleando técnicas fundamentales de aumento de datos (Fundamental Data Augmentation, FDA); y el tercero con la estrategia de aprendizaje para aumentar (Learning-to-Augment Strategy, LAS) y con modelos preentrenados (GoogleNet, ResNet18 y ShuffleNet) alcanzaron exactitudes del 96.88%, 97.50% y



**Fig. 7.** Matriz de confusión normalizada de validación de LeNet5 con tratamiento 11.



**Fig. 8.** Matriz de confusión normalizada de validación de LeNet5 con tratamiento 11.

98.85%, precisiones promedios del 96.93%, 97.73%, 98.81% sensibilidades promedios del 96.87%, 97.5% y 98.75% y especificidades promedios del 98.96%, 99.17% y 99.58% respectivamente con GoogleNet.

En cuanto a la escalabilidad, el método propuesto es una buena alternativa para ser implementado en aplicaciones más amplias de clasificación de madurez de frutas. El uso de las CNN permite procesar grandes volúmenes de datos, lo que es adecuado para sistemas de clasificación en tiempo real en entornos agrícolas. Este modelo mostró tener una alta eficiencia (exactitud de 90.6%). Además, debido a su arquitectura flexible, el modelo puede ser adaptado fácilmente para clasificar diferentes tipos de frutas ajustando mínimamente los parámetros de entrenamiento, lo que sugiere su aplicabilidad en una variedad de contextos agrícolas. Sin embargo, se debe considerar probar este sistema en entornos agrícolas reales, ya que puede ayudar a identificar y corregir posibles deficiencias en el rendimiento del modelo en situaciones prácticas.

La implementación de sistemas de clasificación de madurez de frutas basados en CNN ofrece importantes beneficios para la industria agrícola y alimentaria.

Dicha clasificación optimiza las condiciones de almacenamiento, reduciendo la pérdida de calidad durante este proceso y prolongando la vida útil de las frutas. Esto también facilita una mejor planificación logística en la distribución, asegurando que las fresas lleguen frescas a los puntos de venta y mejorando la satisfacción del cliente mediante una calidad uniforme del producto.

Los datos generados en este estudio, proporciona información valiosa para la toma de decisiones a lo largo de toda la cadena de suministro. La tecnología basada en CNN también es altamente adaptable y escalable, permitiendo su aplicación en diversas áreas agrícolas y geográficas. En conjunto, estos avances tecnológicos pueden transformar significativamente la eficiencia y la calidad en la industria agrícola.

Finalmente, el despliegue de sistemas de inteligencia artificial (IA) en la agricultura y las cadenas de suministro de alimentos tiene importantes implicaciones éticas y sociales. Los beneficios incluyen una mayor eficiencia y productividad, mejor calidad y seguridad alimentaria, reducción de costos y datos más precisos para la toma de decisiones. Es importante abordar estos desafíos mediante políticas y regulaciones adecuadas, en donde se considere la capacitación para agricultores y mejoras continuas.

## **5. Conclusión**

Se determinó que la red neuronal con arquitectura LeNet5, entrenada desde cero con una tasa de aprendizaje de 0.001 y el optimizador RMSProp, logra clasificar correctamente fresas inmaduras, maduras y muy maduras con una exactitud del 90.6%. Además, obtuvo los mejores promedios en las métricas de desempeño: 92.3% en precisión, 90.7% en sensibilidad y 91.7% en puntuación F1.

Aunque estos resultados son prometedores, existen varias oportunidades de mejora.

Una de las más importantes es aumentar el conjunto de datos de imágenes para el entrenamiento, validación y prueba, y asegurar un equilibrio en el número de imágenes para las diferentes clases de madurez de la fruta. También es necesario diversificar las condiciones de las imágenes de prueba, obteniéndolas en diferentes condiciones ambientales y considerando factores no observados previamente, como el desenfoque o la iluminación.

## **Referencias**

1. Acero, L.A., Ong, J.D., Shi, C.J., Dadios E.P.: Strawberry Quality Classification Utilizing Convolutional Neural Network. In: IEEE 13th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), pp. 1–4 (2021) doi: 10.1109/HNICEM54116.2021.9731924.
2. Azizi, H., Asli-Ardeh, E.A., Jahanbakhshi, A.: Vision-based Strawberry Classification Using Generalized and Robust Deep Networks. Journal of Agriculture and Food Research, 15(100931) (2024) doi: 10.1016/j.jafr.2023.100931.
3. Choi, J.Y., Seo, K., Cho, J.S.: Applying Convolutional Neural Networks to Assess the External Quality of Strawberries. Journal of Food Composition and Analysis, 102(104071) (2021) doi: 10.1016/j.jfca.2021.104071.

4. Instituto Tecnológico Superior de Coalcomán: Manual de Producción de fresa en Coalcomán Michoacán (2018)
5. Secretaría de Economía: NMX-FF-062-SCFI-2002. Productos alimenticios no industrializados para consumo humano - fruta fresca - fresa (*Fragaria x ananassa*, dutch) – especificaciones y método de prueba (2002) <http://www.economia-nmx.gob.mx/normas/nmx/2002/nmx-ff-062-scfi-2002.pdf>.
6. Du, X., Bai, J., Plotto, A.: Electronic Nose for Detecting Strawberry Fruit Maturity. *Proc. Fla. State Hort. Soc.*, 123, pp. 259–263 (2010)
7. FAOSTAT: Base de datos estadísticos de la FAO (2022) <http://www.fao.org/faostat/es/#data>.
8. Ganesh, P., Volle, K., Burks, T.F.: Deep Orange: Mask R-CNN Based Orange Detection and Segmentation. *IFAC-PapersOnline*, 52(30), pp. 70–75 (2019) doi: 10.1016/j.ifacol.2019.12.499.
9. Géron, A.: *Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly (2019)
10. Islam, M.R., Matin, A.: Detection of COVID 19 from CT Image by the Novel LeNet-5 CNN Architecture. In: 23rd International Conference on Computer and Information Technology (ICCIT), pp. 1–5 (2020) doi: 10.1109/ICCIT51783.2020.9392723.
11. LeCun, Y., Bottou, L., Bengio, Y.: Gradient-based Learning Applied to Document Recognition. *Proceedings of the IEEE*, 86(11), pp. 2278–2324 (1998) doi: 10.1109/5.726791.
12. Shao, Y., Wang, Y., Xuan, G.: Assessment of Strawberry Ripeness Using Hyperspectral Imaging. *Analytical Letters*, 54(10), pp. 1547–1560 (2020) doi: 10.1080/00032719.2020.1812622.
13. SIACON: Sistema de información agroalimentaria de consulta. Modulo agrícola estatal del SIACON-NG. México: SIAP-SADER (2022) <https://www.gob.mx/siap/documentos/siacon-ng-161430>.
14. Simonyan, K., Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition. *ARXiv* 1409.1556 (2014)
15. Torres, J.: *Python Deep Learning: Introducción práctica con Keras y TensorFlow 2*. Alpha Editorial (2020)
16. Wang, Z., Jin, L., Wang, S.: Apple Stem/Calyx Realtime Recognition Using YOLO-V5 Algorithm for Fruit Automatic Loading System. *Postharvest Biology and Technology*, 185, pp. 111808 (2022) doi: 10.1016/j.postharvbio.2021.111808.
17. Yang, H., Ni, J., Gao, J.: A Novel Method for Peanut Variety Identification and Classification by Improved VGG16. *Scientific Reports*, 11(1) pp. 15756 (2021)
18. Yue, X.Q., Shang, Z.Y., Yang, J.Y.: A Smart Data-Driven Rapid Method to Recognize the Strawberry Maturity. *Information Processing in Agriculture*, 7(4), pp. 575–584 (2020) doi: 10.1016/j.inpa.2019.10.005.
19. Zhou, X., Ampatzidis, Y., Lee, W.S.: Deep Learning-Based Postharvest Strawberry Bruise Detection Under UV and Incandescent Light. *Computers and Electronics in Agriculture*, 202, pp. 107389 (2022) doi: 10.1016/j.compag.2022.107389.
20. Zhou, X., Lee, X.S., Ampatzidis, Y.: Strawberry Maturity Classification from UAV and Near-Ground Imaging Using Deep Learning. *Smart Agricultural Technology*, 1 (2021) doi: 10.1016/j.atech.2021.100001.