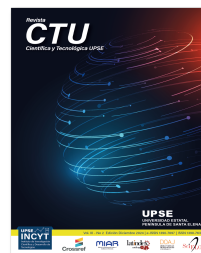


Aprendizaje Profundo aplicado a la clasificación de granos de cacao (*Theobroma cacao L.*) según la calidad de fermentación

Deep Learning Applied to the Classification of Cocoa Beans (*Theobroma cacao L.*) According to Fermentation Quality



Ariosto Eugenio Vicuña Pino¹
Jorge Farouk Molina Noboa¹
Luis Miguel Espín Carrasco¹
Orlando Ramiro Erazo Moreta¹

✉ <https://orcid.org/0000-0002-1459-1759>
✉ <https://orcid.org/0000-0001-8494-5420>
✉ <https://orcid.org/0000-0003-3446-2716>
✉ <https://orcid.org/0000-0001-5642-9920>

¹Universidad Técnica Estatal de Quevedo UTEQ | Quevedo, Los Ríos - Ecuador | CP 120304

✉ avicuna@uteq.edu.ec

<http://doi.org/10.26423/rctu.v11i2.838>
Páginas: 92- 104

Resumen

La fermentación del grano de *Theobroma cacao L.* es un proceso postcosecha importante para el desarrollo de sus propiedades y del aroma. Aunque se trata de un proceso complejo, los agricultores utilizan métodos empíricos/tradicionales para determinar su grado de fermentación, como es la “Prueba de corte” realizada manualmente. No obstante, este tipo de técnicas podría tener una alternativa basada en computadores. Por ello, en este estudio se analizó el uso de redes neuronales convolucionales (CNN) basadas en el aprendizaje profundo, para determinar el grado de fermentación de los granos de cacao. Para el efecto, se desarrolló un modelo cuyo rendimiento fue comprobado en términos de precisión y matriz de confusión. Este modelo alcanzó una exactitud positiva del 82 % y una matriz de confusión con números favorables en los elementos diagonales. Estos resultados dan muestra de que CNN es una opción viable para la clasificación de granos de cacao, según su fermentación.

Palabras clave: Red Neuronal Convolucional, Inteligencia Artificial, Fermentación de cacao, Granos de cacao.

Abstract

The *Theobroma cacao L.* bean fermentation is an important post-harvest process for the development of its properties and aroma. Although cocoa fermentation is complex, farmers use empirical methods to determine its degree of fermentation. One of the traditional techniques used to recognize the quality of fermentation is the “Cut Test”, performed by a person manually. However, this type of techniques could have a computer-based alternative. Therefore, in this study, the use of convolutional neural networks (CNN) based on deep learning was analyzed to determine the degree of fermentation of cocoa beans. For this purpose, a model was developed whose performance was verified in terms of precision and confusion matrix. This model achieved a positive accuracy of 82 % and a confusion matrix with favorable numbers on the diagonal elements. These results show that CNN is a viable option for the classification of cocoa beans based on their fermentation.

Keywords: Convolutional Neural Network, Artificial Intelligence, Cocoa Fermentation, Cocoa Beans.

Recepción: 06/09/2024 | Aprobación: 25/11/2024 | Publicación: 26/12/2024

1. Introducción

El cacao (*Theobroma cacao L.*) es un producto importante a nivel mundial que requiere una atención adecuada que conduzca a que los productos derivados de él sean de buena calidad. Para ello, la fermentación es un paso crítico ya que permite el desarrollo del sabor de sus granos [1], involucrando a componentes como aminoácidos libres, azúcares y péptidos bajo proteólisis [2]. De hecho, los precursores del sabor del cacao se crean durante el procesamiento primario de los granos, lo que incluye la fermentación y el secado [3]. Esto significa que los procesadores requieren de una fermentación completa de los granos de cacao porque aseguran el desarrollo de los precursores del aroma y del sabor [4]. Por lo tanto, la fermentación de los granos de cacao mejora la calidad final del producto, de lo cual depende su valor en los mercados.

Las técnicas de fermentación van desde almacenar el grano de cacao en sacos y que el proceso de fermentación se cumpla, hasta realizar el proceso completo con maquinaria especializada. Cada una de las técnicas da como resultado un grano con diversa calidad de fermentación, razón por la cual, reconocer de manera adecuada la calidad de fermentación es importante para la comercialización de los granos de cacao [4]. Una de las técnicas tradicionales que se usa para reconocer la calidad de fermentación es la “Prueba de corte”, que se la realiza de forma manual y una persona determina la calidad de fermentación. El método es simple. Consiste en cortar una selección de granos de cacao longitudinalmente por el centro para controlar y evaluar su calidad [5], pero requiere de mucho tiempo y se considera subjetivo [6]. Por esta razón, es útil disponer de herramientas que permitan determinar la calidad de fermentación de los granos de cacao de manera económica y ampliamente estandarizada, pudiendo encontrar alguna opción en las Ciencias de la Computación y, en particular, en la Inteligencia Artificial (IA).

En general, las investigaciones relacionadas con la IA en el ámbito de la agricultura se han popularizado en años recientes. Para que ello sea posible, ha sido necesario recurrir a técnicas como el aprendizaje automático (ML, Machine Learning) y el Aprendizaje profundo (DL, Deep Learning). El ML es capaz de producir un comportamiento utilizando algoritmos que a su vez son alimentados por una gran cantidad de datos [7]. Por su parte, el DL es un subconjunto de ML que se utiliza para comprender conceptos de mayor precisión, como la creación de clasificadores robustos y la elección de técnicas de extracción de características que debe medirse cuidadosamente para proporcionar información representativa al clasificador [8]. A esto se le puede agregar las redes neuronales convolucionales (CNN, Convolutional Neural Network) que han demostrado ser útiles para la detección de objetos y la clasificación de imágenes [9, 10, 11, 12]. Estas pueden ser aplicadas en la clasificación de imágenes de manera más eficiente [13], logrando además aprovecharlas en otras tareas que requieren menos datos de entrenamiento [14].

A partir de este tipo de técnicas de IA se han llevado a cabo varios trabajos para analizar imágenes en el ámbito agrícola [15, 16], llegando incluso a la posibilidad de suplantar a los métodos tradicionales como la “Prueba de corte” [17]. Una primera opción es emplear CNN para extraer características de los granos de cacao. Para ello, Adhitya *et al.* [8] utilizaron

imágenes de granos de cacao de diferentes niveles de calidad, logrando, con cinco capas de CNN, clasificar siete tipos de imágenes con una tasa de precisión del 82%. Otra opción es usar la visión por computadora como un método rápido y preciso para clasificar los granos de cacao fermentados. En este caso, se dividieron los granos seleccionados en cuatro etapas de fermentación empleando bosques aleatorios y alcanzando una precisión superior al 80% [1]. Otros trabajos más recientes se han limitado a clasificar los granos de cacao como saludables o enfermos [18]. En definitiva, estos trabajos son una muestra de la utilidad de la IA aplicada al cacao, pero es necesario también analizar su idoneidad con muestras e imágenes de cacao de otras regiones, como en el caso del cacao ecuatoriano.

Teniendo en cuenta este contexto y centrándose en la agricultura ecuatoriana, surge la pregunta: ¿Cómo identificar la calidad de fermentación en los granos de cacao ecuatoriano? Para responderla, este trabajo se basa en el uso de IA, y en particular de CNN; es decir, se tiene como objetivo diseñar un modelo de clasificación de imágenes empleando una técnica de DL para medir la calidad de fermentación de los granos de cacao ecuatoriano y seleccionar la mejor opción para reconocer patrones a partir de la información digital. Con esto se busca llegar a una identificación de la calidad de la fermentación de los granos de cacao más precisa, en comparación con las técnicas tradicionales. Así, se aspira a aportar en la selección de la técnica que permita mejorar los tiempos al momento de reconocer la calidad de fermentación y para que este no vea afectado su precio en los mercados.

2. Materiales y Métodos

En términos generales, el estudio se basó en el uso del algoritmo CNN de *Deep Learning* dedicado al reconocimiento de imágenes, enfocado a la calidad de fermentación de los granos de cacao. CNN es una red neuronal artificial que imita el funcionamiento del ojo humano para identificar objetos o características, fragmenta las imágenes y se entrena usando capas especializadas [19, 20]. El algoritmo obtiene las características sobre la calidad de fermentación desde una base de conocimiento con imágenes clasificadas en los cuatro tipos de calidad de fermentación. Además, se empleó el análisis de texturas GLCM (*Gray Level Co-Occurrence Matrix*) como método estadístico que permite caracterizar la textura de una imagen calculando la frecuencia con la que se producen pares de píxeles con valores específicos en un espacio de la imagen [21]. A continuación, se describen más detalles del estudio.

Conjunto de datos

Las imágenes utilizadas en el estudio fueron obtenidas a partir de granos de cacao recolectados en Quevedo, Ecuador. Los granos de la muestra estaban secos para una obtención adecuada de las imágenes. Estos granos fermentados fueron clasificados manualmente por docentes de Ingeniería Agrícola/Agro-nómica, expertos en el tema, y pertenecientes a la misma institución de los autores, utilizando la prueba de corte. Para ello, se aplicó la norma ecuatoriana INEM 176:2006 [22] a todos los granos de la muestra; es decir, se clasificaron en buena fermentación, mediana fermentación, grano violeta, grano pizarroso y grano mohoso. Sin embargo, el conjunto de datos no dispuso de

imágenes para el tipo de fermentación pizarroso. Así, el conjunto de datos resultante tuvo 212 imágenes, teniendo en cuenta lo realizado en otro trabajo [23].

Procesamiento de datos

Previo al análisis necesario, se efectuó un preprocesamiento de datos. Se leyeron las imágenes de entrada, seguidas de un proceso de reescalado aplicado a todas las imágenes del conjunto de datos [[24]]. Las imágenes, categorizadas en los cuatro tipos de calidad de fermentación citadas, se organizaron en carpetas de archivos para facilitar su lectura y procesamiento desde el software usado.

Preparación del conjunto de datos

Las imágenes recolectadas se usaron para entrenar la red neuronal convolucional, empleando las categorías como los

objetos de experimentación.

Para ello fue necesaria una preparación preliminar debido a que las imágenes a ingresar al modelo formulado deben ser del mismo tamaño, y generalmente no se ingresan imágenes de alta resolución, sino de dimensiones reducidas, puesto que mientras mayor sea el tamaño de la imagen, mayor será el poder computacional requerido y el tiempo de entrenamiento del modelo.

Además, para el tipo de imagen de este trabajo en el que los patrones a detectar son detalles muy pequeños en la imagen, es necesario que la imagen tenga un tamaño considerable.

Así, para mantener un equilibrio, una vez que la imagen tuvo la forma deseada se empleó el método *resize* para redimensionar la imagen a un tamaño de 156 x 256 píxeles, como se observa en la Figura 1.



Figura 1. Imágenes de los granos de cacao por tipo de fermentación.

Construcción del modelo CNN

Aunque existen varios lenguajes de programación para entrenar modelos de clasificación DL, como R, *DeepCC* desarrollado en R y C [25], Matlab [26], o Python [27, 28, 29], en este trabajo se eligió Python.

El motivo radica en que este lenguaje se destaca como uno de los principales lenguajes de programación en la comunidad para implementar marcos de DL [30] y a que cuenta con librerías de código abierto como Keras, TensorFlow, openCV [27, 29], que facilitan el reconocimiento de imágenes sin necesidad de consumir excesivamente recursos computacionales [31, 32]. Así, y con base en la experiencia previa de los autores y recomendaciones de expertos en el tema, se tomó la decisión de usar Python con las librerías Keras y TensorFlow.

Arquitectura del modelo CNN

El modelo de CNN construido está compuesto de tres capas principales: una de entrada, una oculta y una de salida. En la Tabla 1 se pueden visualizar los parámetros de cada capa. Ellos fueron fijados tomando en cuenta las características de las imágenes, lo que hizo necesario recortarlas apuntando a conseguir un mejor resultado en el entrenamiento.

El primer parámetro es un conjunto de **kernels** (llamado filtros); por cada filtro se obtiene una matriz de salida. La primera capa (Entrada) solo lee la imagen de entrada por lo que no hay parámetros por aprender aquí. La segunda capa, **Pooling layers**, toma los mapas de características en la entrada y tiene k mapas de características de salida; son 32 mapas de características como entrada y 64 mapas de características de salida. Los parámetros de las capas *Pooling layers* son utilizados para realizar un *Subsampling con Max-Pooling* (como se explica a detalle más adelante).

Por último, la capa de salida es una capa normalmente conectada, las dimensiones del volumen de salida son $1 \times 1 \times N$, donde N es el número de clases de salida que se están evaluando, teniendo en cuenta que se tiene 256 canales de dimensión. En esta capa hay una conexión entre todas sus neuronas y con cada neurona de la capa anterior.

Además, la planificación de la representación gráfica del modelo, con cada uno de sus parámetros se puede visualizar en la Figura 2, donde se aprecia la entrada y salida de las neuronas utilizadas en las capas de entrada, ocultas y de salida de los canales de convolución de la red neuronal convolucional.

Tabla 1: Parámetros CNN de 2 capas (2CNN)

Capa	Tipo	Parámetros	Descripción
Capa 1	Entrada	-	Esta capa recibe la imagen de entrada y la procesa sin parámetros a aprender. Se encarga de la lectura inicial de los datos de imagen.
Capa 2	Pooling Layers	- Entrada: 32 mapas de características - Salida: 64 mapas de características	Realiza un proceso de Subsampling utilizando Max-Pooling. Esta técnica reduce la dimensionalidad de los mapas de características, preservando la información más relevante. La capa de pooling ayuda a mejorar la robustez del modelo ante variaciones en la entrada.
Capa 3	Capa de Salida	- Dimensiones de salida: 1x1xN - N: número de clases de salida - Canales: 256	Es una capa totalmente conectada, donde cada neurona está conectada a todas las neuronas de la capa anterior. La dimensión de salida 1x1xN indica que produce un único valor para cada clase a evaluar. Se utiliza para clasificar la entrada en las diferentes categorías.

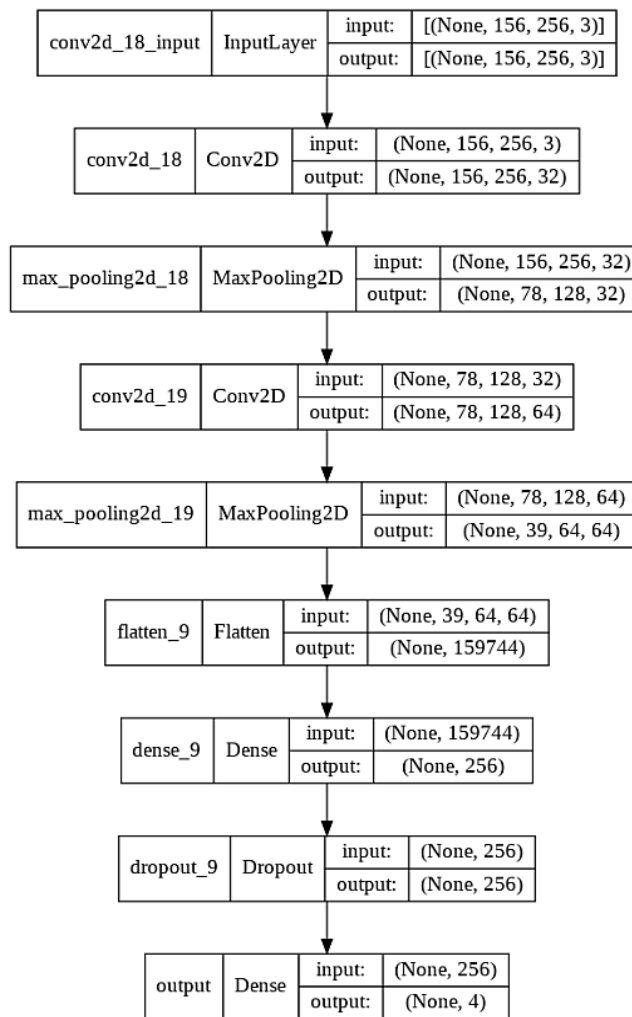


Figura 2. Planificación de la representación gráfica del modelo.

Las imágenes utilizadas como entrada del modelo tienen una dimensión de 156 x 256 píxeles (ancho y alto) con dos canales. Así también, la función de activación utilizada para este modelo es ReLu, la cual se usa tanto en el primero como en el segundo filtro convolucional. Las capas están compuestas por neuronas y estas se conectan a una pequeña

región de la capa anterior, lo que quiere decir que no están completamente conectadas [33]. El primer canal utiliza 32 filtros convolucionales de tamaño 3x3, con una resolución del tamaño de un factor 2 y esa capa es llamada MaxPool2D (ver Figura 3).

Primera Convolución

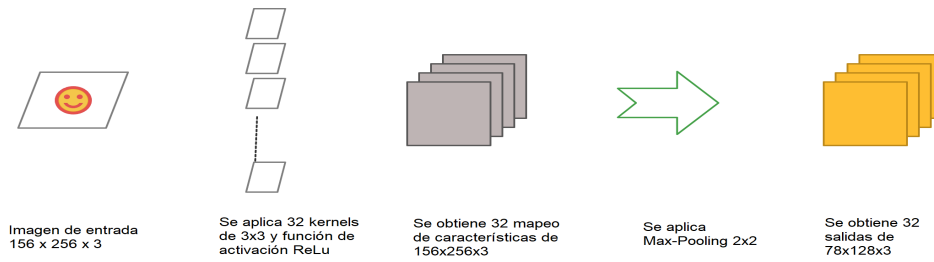


Figura 3. Primera convolución.

En el segundo canal se utilizaron 64 filtros convolucionales de tamaño 2x2 (ver Figura 4), con una resolución del tamaño de un factor 2 igual al primer filtro, puesto que dieron los mejores resultados en el entrenamiento. Fue necesario además añadir una nueva capa, llamada “Flatten”. Esta capa permite unir la zona de la capa convolucional con la capa

fully connected; todas las neuronas de una capa oculta están conectadas a todas las neuronas de la capa que les precede y sigue [34]. Así, una imagen muy profunda/pequeña pasará a ser plana con la ayuda de esta capa, pasando la información de la red neuronal a estar en una sola dimensión.

Segunda Convolución

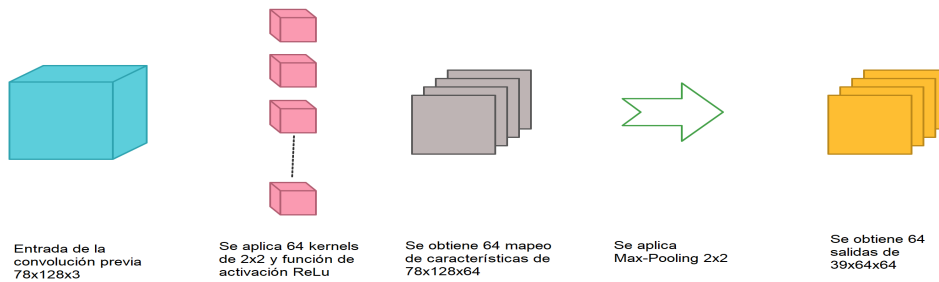


Figura 4. Segunda convolución.

En la capa fully connected (Figura 5) se encuentran las capas densas. Después de aplanar la información, es recibida por una capa normal; es decir, ahora se añade una capa donde todas las neuronas estarán conectadas con las neuronas de la capa pasada. En esta capa serán 256 neuronas. Para evitar el

overfitting (que no estén conectadas al 100 % de las neuronas de una capa con respecto a la otra), es necesario utilizar la capa Dropout. En otras palabras, se apagará el 50 % de las neuronas cada paso para evitar llegar a sobre ajustar.

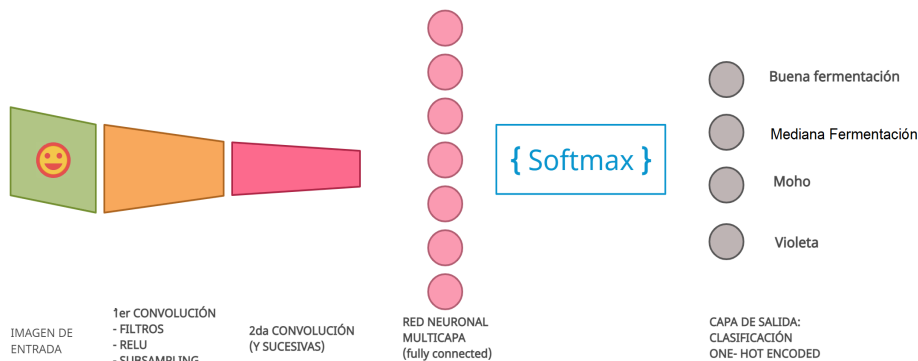


Figura 5. Arquitectura CNN.

Por último, en la capa de salida es necesario tener la función de activación en “*softmax*” para que genere el resultado como si fueran probabilidades. Para ello se cuenta con cuatro clases

o categorías diferentes y es necesario que la capa de salida tenga el tamaño de ellas. Así, la Figura 6 muestra el modelo de la red neuronal convolucional.

```

Model: "sequential"
-----
Layer (type)                Output Shape              Param #
-----
conv2d (Conv2D)             (None, 156, 256, 32)     896
max_pooling2d (MaxPooling2D) (None, 78, 128, 32)      0
conv2d_1 (Conv2D)           (None, 78, 128, 64)     8256
max_pooling2d_1 (MaxPooling2D) (None, 39, 64, 64)      0
flatten (Flatten)           (None, 159744)           0
dense (Dense)                (None, 256)              40894720
dropout (Dropout)           (None, 256)              0
output (Dense)              (None, 4)                1028
-----
Total params: 40,904,900
Trainable params: 40,904,900
Non-trainable params: 0
    
```

Figura 6. Modelo de la red neuronal convolucional.

Descripción del modelo CNN

Se usaron métricas para monitorear el proceso de aprendizaje (y prueba) de la red neuronal. En este modelo, solo se consideró la *accuracy* (fracción de imágenes que son correctamente clasificadas). La función de *loss* especificada fue *categorical_crossentropy*. La función de *loss* es uno de los parámetros necesarios para cuantificar qué tan cerca está una determinada red neuronal de su estado ideal durante el proceso de entrenamiento. El optimizador utilizado fue el *Adaptive moment estimation* (Adam), que es una combinación de las bondades de otros dos algoritmos: *AdaGrad* y *RMSProp* [35]. La métrica fue *accuracy*; con ella se evalúa el porcentaje de aciertos averiguando dónde el modelo predice la etiqueta correcta.

Estructura del entrenamiento del modelo CNN

Para esta parte, se utilizó la función *modelcnn.fit_generator()*, que recibe como argumentos las variables *train_generator()* y *validation_generator()* asignadas a las funciones:

training_datagen.flow_from_directory() y,

test_datagen.flow_from_directory(),

utilizadas en la creación del modelo. De igual forma, se requiere especificar la cantidad de épocas o veces que se va a repetir el entrenamiento; esa cantidad se encuentra almacenada en la variable *epochs*, que en este caso fue 80 épocas. Los datos de la validación también son necesarios para el entrenamiento del modelo y estos se encuentran en la variable *val_generator*.

A partir del modelo CNN entrenado en el entorno gratuito de *Jupyter Notebook*, que no requiere configuración y que se ejecuta completamente en la nube (*Google Colab*) [36], se realizaron varias pruebas para comprobar que este funciona apropiadamente. Con el modelo ya entrenado se realizó la predicción en *Google Colab* utilizando imágenes de prueba. También, hay que tener presente que el modelo se entrenó con el algoritmo de Descenso de gradientes estocástico (*Stochastic Gradient Descent* - SGD) [37], aplicando cuatro ensayos de entrenamiento seguidos de 50, 60, 70 y 80 épocas, con una tasa de aprendizaje de 0,0001 del algoritmo de optimización Adam [24] y un minilote [38] de 32. Así, la Tabla 2 detalla las especificaciones de *Google Colab* usadas.

Tabla 2: Especificaciones de *Google Colab* para el estudio

CPU	1 x procesador Xeon de un núcleo con hiperproceso a 2,3 Ghz (1 núcleo, 2 subprocesos)
RAM	12Gb GDDR5 VRAM
Disco	33 Gb
GPU	1xTesla K80, cómputo 3.7, con 2496 núcleos CUDA
Python	3
Tensorflow	2.8.0
SO	macOS Mojave 10.14.2

3. Resultados

Clasificación

Las tasas de clasificación obtenidas para diferentes números de épocas se presentan en la Tabla 3. Se puede observar que

se mantiene el valor de la tasa de clasificación al incrementar el número de épocas, lo cual podría deberse a la cantidad de datos utilizados para el entrenamiento de los modelos CNN. De esta manera, se obtuvo una precisión máxima de 1,0000 para 80 épocas de entrenamiento.

Tabla 3: Resultados de clasificación

Épocas	Número de iteraciones	Tiempo de entrenamiento	Efectividad - Acuracy (%)	Pérdida - Loss (%)
50	6	609 ms/step	1,0000	0,0075
60	6	596 ms/step	1,0000	0,0061
70	6	635 ms/step	1,0000	1,3684e-04
80	6	626 ms/step	1,0000	0,0026

Predicción

Los resultados de la predicción, con base en la calidad de la fermentación, pueden dividirse en dos partes: (1) una evaluación de la robustez y la viabilidad del modelo, y (2) una evaluación de la precisión de la clasificación de granos de cacao [39]. En cuanto a lo primero, la Figura 7 muestra las variaciones en la precisión y pérdida del entrenamiento

durante la época de prueba basada en CNN. Se puede ver que después de 80 épocas, las precisiones de CNN son efectivas. Además, la curva de precisión es efectiva en comparación con la de pérdida [39]. Adicionalmente, la Figura 8 muestra las variaciones en la precisión y pérdida de los datos de validación. Las precisiones después del entrenamiento con 80 épocas son efectivas. Además, la curva de precisión es efectiva en comparación con la pérdida [39, 40].

Training Loss and Accuracy on Dataset - Regarding Training

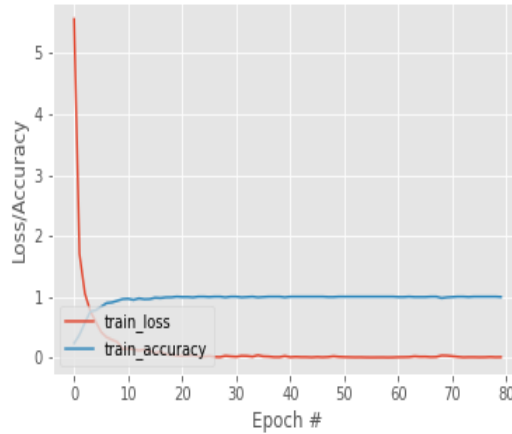


Figura 7. Pérdida de entrenamiento y precisión en el conjunto de datos con 80 épocas - En relación con los datos de entrenamiento.

Training Loss and Accuracy on Dataset - Regarding validatio

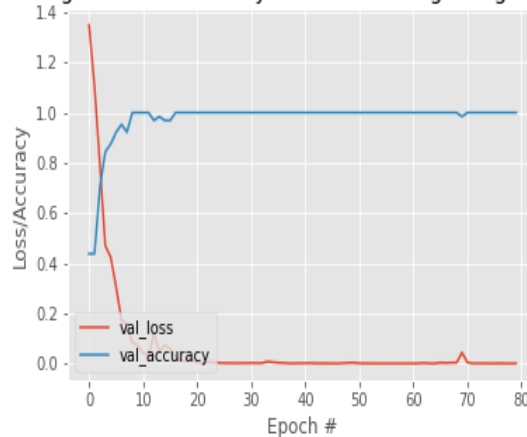


Figura 8. Pérdida de entrenamiento y precisión en el conjunto de datos con 80 épocas - En relación con los datos de validación

Por otra parte, la Figura 9 muestra las variaciones en la precisión y pérdida tanto de los datos de entrenamiento como de los de validación. Las precisiones después del

entrenamiento con 80 épocas son efectivas para los datos de entrenamiento y validación. Las curvas de precisiones son efectivas comparadas con las curvas de pérdidas [39, 40].



Figura 9. Pérdida de entrenamiento y precisión en el conjunto de datos con 80 épocas - En relación con los datos de validación y entrenamiento

Los resultados de la evaluación de la precisión de la clasificación también pueden dividirse en dos partes. La primera se resume en una matriz de confusión de la muestra, donde se refleja el número de muestras en cada combinación de clase de referencia y clase predicha [41]. La segunda corresponde a las métricas de clasificación del modelo CNN.

La matriz de confusión basada en CNN se muestra en la Figura 10. Cuanto mayor sea el número de elementos diagonales, mejor será la clasificación [39], mientras los

otros muestran el número de clasificaciones inexactas [42]. Haciendo uso de las mismas imágenes de entrenamiento, cada una de las imágenes del conjunto de datos es utilizado como datos de prueba para verificar qué tan efectivo es el entrenamiento. De esta misma figura se desprende que las muestras de buena fermentación se dividen a menudo en las otras clases por error. Esto puede estar relacionado a la cantidad de imágenes del entrenamiento y a que los datos de buena fermentación son muy parecidos a los datos de las otras clases, similar a lo ocurrido en trabajo relacionado [43].

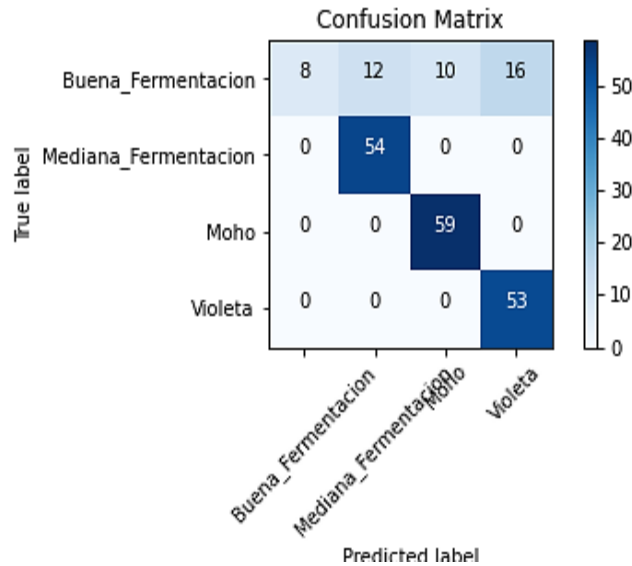


Figura 10. Matriz de confusión.

Por otro lado, (Figura 11) se muestra el detalle de las diferentes métricas de clasificación que se usaron para determinar la exactitud del modelo CNN. Se emplearon las métricas de *precision*, *recall* y *F1* para cada uno de los tipos de calidad de fermentación. La exactitud para este modelo se ve reflejada en el *Accuracy*, con un porcentaje de 82% de exactitud del modelo CNN [44]. Debe notarse también

que este valor es comparable con los obtenidos en otros estudios relacionados a cacao como los [8] y [1]. En el primer caso, los autores también reportaron un *Accuracy* de 82% utilizando cinco capas, mientras en este trabajo se emplearon solo dos. En el segundo, el valor de *Accuracy* fue mayor al aquí obtenido gracias al mayor conjunto de imágenes, pero con valores aproximadamente iguales para *precision* y *recall*

	precision	recall	f1-score	support
0	1.0000	0.1739	0.2963	46
1	0.8182	1.0000	0.9000	54
2	0.8551	1.0000	0.9219	59
3	0.7681	1.0000	0.8689	53
accuracy			0.8208	212
macro avg	0.8603	0.7935	0.7468	212
weighted avg	0.8554	0.8208	0.7673	212

Figura 11. Reporte de las métricas de clasificación.

4. Discusión

En este trabajo se presentan resultados prometedores para la clasificación automática de granos de cacao, pero hay algunos aspectos sobre los cuales se puede reflexionar. En primer lugar, durante el entrenamiento del modelo CNN se presentaron ciertos errores en la clasificación de las muestras. Esto puede atribuirse al tamaño del conjunto de datos utilizado. La precisión de la visión por ordenador fue excelente a pesar de haber tenido que recortar las imágenes. Después de 80 épocas, el modelo alcanzó una precisión del 82 %, lo cual es aceptable para un conjunto de datos dividido en 4 categorías diferentes. Es de esperar que si se incrementa la cantidad de imágenes (por ejemplo, a mil por categoría), el modelo aprenda las características distintivas de cada categoría, resultando en una clasificación más precisa y una menor confusión entre categorías similares.

En segundo lugar, se podría considerar un incremento en el tamaño del conjunto de datos. Se pueden considerar dos estrategias para abordar en un futuro estudio. La primera opción es incrementar el tamaño del conjunto de datos mediante técnicas de aumento de datos, como rotación, recorte, cambio de iluminación, o adición de ruido [45], lo que permitiría crear variaciones adicionales a partir de las imágenes existentes. Aunque esto podría incrementar el volumen del conjunto de datos derivando en que el modelo generalice mejor las variaciones en las imágenes reales, debe ser verificado con más detenimiento. Esto conduce a pensar en la segunda opción, la captura de imágenes en un entorno controlado. Esta captura puede contribuir a que las características relevantes del grano de cacao sean consistentes entre las diferentes categorías. En definitiva, disponer de más imágenes por categoría, idealmente mil o más, permitirá que el modelo aprenda de manera más efectiva logrando una clasificación más precisa.

En tercer lugar, y a pesar de la limitación antes mencionada, los resultados obtenidos son comparables con otros trabajos. De partida, la precisión aquí reportada es aproximada a la alcanzada por otros autores que emplearon imágenes de cacao de otros países (80 % [1] y 82 % [8]), pero usando conjuntos de datos con mayor cantidad de imágenes. También se puede comparar el rendimiento de CNN, pero con estudios que clasifican otros tipos de granos. Por ejemplo, un estudio centrado en clasificar granos de trigo alcanzó una precisión del 94.88 % en el entrenamiento y un 97.53 % en la prueba, logrando reducir las pérdidas al 15 % en el entrenamiento y al 8 % en la prueba, pero gracias a un conjunto de 15000 imágenes [46]. Otro trabajo, pero dedicado a la clasificación de granos de café [47], preparó 72000 imágenes mediante calibración de datos, con 36000 de granos buenos y 36000

de granos malos. Seleccionando al azar 7000 imágenes, los autores reportaron una precisión del 94.63 %. Aunque estos valores son superiores a los reportados en este trabajo, hay que rescatar que la precisión lograda es prometedora en comparación con los otros tamaños de los conjuntos de datos.

Finalmente, el modelo puede evolucionar hasta llegar a ser incorporado en una aplicación móvil que facilite el proceso de clasificación. La viabilidad práctica de esta idea está siendo explorada por los autores. Para el efecto, se ha considerado la capacidad de *TensorFlow* para convertir el modelo a *TensorFlow Lite*. Esta conversión permite integrarlo sin problemas en una aplicación desarrollada en Android Studio, lo cual es especialmente beneficioso, ya que el modelo entrenado puede funcionar sin necesidad de conexión a Internet. Esto significa que los usuarios pueden acceder a las funcionalidades de la aplicación incluso en áreas remotas donde la señal telefónica es escasa o inexistente. Al eliminar la dependencia de la conectividad, se amplía el rango de aplicaciones posibles, permitiendo que la aplicación se utilice en contextos diversos y reales, como en zonas rurales o durante actividades al aire libre, lo que maximiza su utilidad y accesibilidad.

5. Conclusiones

Partiendo de la importancia de la clasificación de granos de cacao basada en su calidad de fermentación, en este trabajo se ha abordado la construcción y uso de un modelo CNN de aprendizaje profundo. El aprendizaje profundo, a través del análisis de imágenes, ofrece una opción viable para este tipo de clasificación. Esto se ha evidenciado a partir del desarrollo de un modelo cuyo rendimiento es comparable al de estudios similares, pero con datos de una zona diferente. La exactitud positiva alcanzada por el modelo fue del 82 % empleando un entrenamiento de 80 épocas.

Además, la matriz de confusión reveló que las muestras de mediana fermentación, moho y violeta mantienen una buena clasificación con el número máximo de muestras en los elementos diagonales. Sin embargo, varias de las muestras se dividen a menudo por error en otras clases, lo que podría deberse a que tales muestras son muy similares, teniendo en cuenta la cantidad de datos usada para el entrenamiento. Este aspecto, o en general la precisión del modelo, podría mejorarse al efectuar el entrenamiento con un conjunto de datos mucho más amplio. Independiente de ello, el modelo obtenido podría ser empleado para la construcción de una herramienta de software destinada a profesionales o agricultores interesados en clasificar granos de cacao por su fermentación utilizando, por ejemplo, un teléfono inteligente.

En el futuro se espera abordar estas ideas como una manera de aportar a la sostenibilidad de la industria del grano del cacao, a los agricultores, a los procesadores y en general a quienes dependen de esta fuente de ingresos.

Financiamiento:

Los autores expresan que la financiación fue llevada por la Universidad Técnica Estatal de Quevedo (Quevedo, Ecuador), proyecto FOCICYT (PFOC7-35-2020), Séptima Convocatoria. Los autores agradecen a Cristian Zambrano por facilitar el conjunto de datos para el estudio.

Conflicto de intereses:

Los autores declaran no tener conflicto de intereses.

Contribución de autor/es:

Bajo los lineamientos CRediT (Taxonomía de Roles de Contribuyente), los autores dan fe de las contribuciones realizadas al trabajo de investigación, que se detallan: Ariosto Vicuña: revisión, redacción, validación, supervisión, administración del proyecto, adquisición de financiación, recursos, análisis, conducción, metodología, conceptualización. Jorge Molina: visualización, redacción, validación, administración del proyecto, recursos, software, análisis, conducción, curación de datos, conceptualización. Luis Espín: visualización, validación, software, análisis, conducción, curación de datos, conceptualización. Orlando Erazo: visualización, revisión, redacción, supervisión, administración del proyecto, recursos, conceptualización.

6. Referencias

- OLIVEIRA, MARCIANO M., CERQUEIRA, BRENO V., BARBON, SYLVIO AND BARBIN, DOUGLAS F. Classification of fermented cocoa beans (cut test) using computer vision. *Journal of Food Composition and Analysis* [online]. 2021, vol. 97, pág. 103771. Disponible en: DOI%2010.1016/J.JFCA.2020.103771.
- HUE, C., GUNATA, Z., BREYSSE, A., DAVRIEUX, F., BOULANGER, R. AND SAUVAGE, F. X. Impact of fermentation on nitrogenous compounds of cocoa beans (*Theobroma cacao* L.) from various origins. *Food Chemistry* [online]. 2016, vol. 192, págs. 958-964. Disponible en: DOI%2010.1016/J.FOODCHEM.2015.07.115.
- CAMU, NICHOLAS, DE WINTER, TOM, ADDO, SOLOMON K., TAKRAMA, JEMMY S., BERNAERT, HERWIG AND DE VUYST, LUC. Fermentation of cocoa beans: influence of microbial activities and polyphenol concentrations on the flavour of chocolate. *Journal of the Science of Food and Agriculture* [online]. 2008, vol. 88, n.º 13, págs. 2288-2297. Disponible en: DOI%2010.1002/JSFA.3349.
- MORÁN, BAJAÑA, ANGEL, LUIS, GALLARDO, PALADINES AND AARÓN, JOSUÉ. Diseño e implementación de un prototipo clasificador de granos de cacao usando Visión Artificial y Machine Learning. *Journal of the Science of Food and Agriculture* [online]. 2021. Disponible en: <http://dspace.upse.edu.ec/handle/123456789/21405>.
- VAN DURME, JIM, INGELS, ISABEL AND DE WINNE, ANN. Inline roasting hyphenated with gas chromatography–mass spectrometry as an innovative approach for assessment of cocoa fermentation quality and aroma formation potential. *Food Chemistry* [online]. 2016, vol. 205, págs. 66-72. Disponible en: DOI%2010.1016/J.FOODCHEM.2016.03.004.
- KONGOR, JE, TAKRAMA, JF, BUDU, AS AND MENSAH-BROWN, H. Effects of fermentation and drying on the fermentation index and cut test of pulp pre-conditioned Ghanaian cocoa (*Theobroma cacao*) beans. [Online]. 2013. Disponible en: <http://csirspace.csirgh.com/handle/123456789/528>.
- EL NAQA, ISSAM AND MURPHY, MARTIN J. What Is Machine Learning? *Machine Learning in Radiation Oncology* [online]. 2015, págs. 3-11. Disponible en: DOI%2010.1007/978-3-319-18305-3_1.
- ADHITYA, YUDHI, PRAKOSA, SETYA WIDYAWAN, KÖPPEN, MARIO AND LEU, JENQ SHIOU. Convolutional Neural Network Application in Smart Farming. *Springer* [online]. 2019, vol. 1100, págs. 287-297. Disponible en: DOI%2010.1007/978-981-15-0399-3_23.
- ZHOU, BOLEI, KHOSLA, ADITYA, LAPEDRIZA, AGATA, OLIVA, AUDE AND TORRALBA, ANTONIO. Object Detectors Emerge in Deep Scene CNNs. *Computer Vision and Pattern Recognition (cs.CV)* [online]. 2014. Disponible en: <https://arxiv.org/abs/1412.6856v2>.
- SANTOS, LUÍS, SANTOS, FILIPE N., OLIVEIRA, PAULO MOURA AND SHINDE, PRANJALI. Deep Learning Applications in Agriculture: A Short Review. *Advances in Intelligent Systems and Computing* [online]. 2019, vol. 1092, págs. 139-151. Disponible en: DOI%2010.1007/978-3-030-35990-4_12.
- SIDDIQUI, ZAHID ALI, PARK, UNSANG, LEE, SANG WOONG, JUNG, NAM JOON, CHOI, MINHEE, LIM, CHANUK AND SEO, JANG HUN. Robust Powerline Equipment Inspection System Based on a Convolutional Neural Network. *Sensors* [online]. 2018, vol. 18,

- n.º 11, pág. 3837. Disponible en: DOI%2010 . 3390/S18113837.
12. SRINIVAS, SURAJ, SARVADEVABHATLA, RAVI KIRAN, MOPURI, KONDA REDDY, PRABHU, NIKITA, KRUTHIVENTI, SRINIVAS S.S. AND BABU, R. VENKATESH. A taxonomy of deep convolutional neural nets for computer vision. *Frontiers Robotics AI* [online]. 2016, pág. 36. Disponible en: DOI%2010 . 3389/ FROBT . 2015 . 00036/BIBTEX.
 13. HOWARD, ANDREW G. Some Improvements on Deep Convolutional Neural Network Based Image Classification. *Computer Vision and Pattern Recognition* [online]. 2013. Disponible en: <https://arxiv.org/abs/1312.5402v1>.
 14. JMOUR, NADIA, ZAYEN, SEHLA AND ABDELKRIM, AFEF. Convolutional neural networks for image classification. *International Conference on Advanced Systems and Electric Technologies* [online]. 2018, págs. 397-402. Disponible en: DOI%2010 . 1109/ASET . 2018 . 8379889.
 15. FERRANCOL FERRAN, ZHAIRA MARIANNE, TAMAYO ONGQUIT, LUISA ANN AND ROMEROSO ARBOLEDA, EDWIN. A review of machine learning applications in Cacao Post-harvest management. *International Journal of Science and Research Archive* [online]. 2024, vol. 2024, n.º 01, págs. 1540-1550. Disponible en: DOI%2010 . 30574/ ijsra . 2024 . 11 . 1 . 0186.
 16. OMAS-AS, ANDREY M., ARBOLEDA, EDWIN R. AND DAANG, JEWELLE ABBYGALE M. Machine Learning as a Strategic Tool: A Comprehensive Literature Review for Advancing Agricultural Analysis, with Emphasis on the Cocoa Bean Quality Assessment. *International Journal of Scientific Research and Engineering Development* [online]. 2024, vol. 7, n.º 01. Disponible en: DOI%2010 . 5281/zenodo . 10552894.
 17. JHA, KIRTAN, DOSHI, AALAP, PATEL, POOJAN AND SHAH, MANAN. A comprehensive review on automation in agriculture using artificial intelligence. *Artificial Intelligence in Agriculture* [online]. 2019, vol. 2, págs. 1-12. Disponible en: DOI%2010 . 1016/J . AIIA . 2019 . 05 . 004.
 18. ZHININ-VERA, LUIS, ZHIMINAICELA-CABRERA²⁵. JONATHAN, PRETEL, ELENA, SUÁREZ, PAMELA, CHANG, OSCAR, CASTRO, FRANCESC ANTÓN AND LÓPEZ DE LA ROSA, FRANCISCO. Artificial Vision Technique to Detect and Classify Cocoa Beans. *Springer Nature Switzerland* [online]. 2019, págs. 217-228. Disponible en: DOI%2010 . 1007/ 978-3-031-43078-7_18.
 19. LECUN, YANN, BOTTOU, LÉON, BENGIO, YOSHUA AND HAFFNER, PATRICK. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* [online]. 1998, vol. 86, n.º 11, págs. 2278-2323. Disponible en: DOI%2010 . 1109/5 . 726791.
 20. GU, JIUXIANG, WANG, ZHENHUA, KUEN, JASON, MA, LIANYANG, SHAHROUDY, AMIR, SHUAI, BING, LIU, TING, WANG, XINGXING, WANG, GANG, CAI, JIANFEI AND CHEN, TSUHAN. Recent advances in convolutional neural networks. *Pattern Recognition* [online]. 2018, vol. 77, n.º 11, págs. 354-377. Disponible en: DOI%2010 . 1016/ J . PATCOG . 2017 . 10 . 013.
 21. HERNÁNDEZ PACHECO, DENIS. Análisis de textura basado en la matriz de ocurrencia de niveles de gris. *Facultad de Matemática, Física y Computación. Departamento de Ciencias de la Computación* [online]. 2008. Disponible en: <http://dspace.uclv.edu.cu:8089/xmlui/handle/123456789/7625>.
 22. INSTITUTO ECUATORIANO DE NORMALIZACIÓN. Cacao en grano. Requisitos. 2006.
 23. HERNÁNDEZ-CALVARIO, OMAR, FLORIÁN, FRIDA, SÁNCHEZ, MARÍA GUADALUPE, ÁVILA-GEORGE, HIMER, HERNÁNDEZ-CALVARIO, OMAR, FLORIÁN, FRIDA, SÁNCHEZ, MARÍA GUADALUPE AND ÁVILA-GEORGE, HIMER. Conteo de plantas de agave usando redes neuronales convolucionales e imágenes adquiridas desde un vehículo aéreo no tripulado. *RISTI - Revista Ibérica de Sistemas e Tecnologías de Informação*. 2022, vol. 2022, n.º 45, págs. 64-76. Disponible en: DOI%2010 . 17013/ RISTI . 45 . 64-76.
 24. ALQUDAH, ALI MOHAMMAD, ALQURAN, HIAM AND QASMIEH, ISAM ABU. Segmented and Non-Segmented Skin Lesions Classification Using Transfer Learning and Adaptive Moment Learning Rate Technique Using Pretrained Convolutional Neural Network. *Journal of Biomimetics, Biomaterials and Biomedical Engineering*. 2019, vol. 42, n.º 45, págs. 67-78. Disponible en: DOI%2010 . 4028/ WWW . SCIENTIF%09IC . NET/ JBBBE . 42 . 67.
 - GAO, FENG, WANG, WEI, TAN, MIAOMIAO, ZHU, LINA, ZHANG, YUCHEN, FESSLER, EVELYN, VERMEULEN, LOUIS AND WANG, XIN. DeepCC: a novel deep learning-based framework for cancer molecular subtype classification. *Oncogenesis*. 2019, vol. 8,

- n.º 9, págs. 1-12. Disponible en: DOI % 2010 . 1038/s41389-019-0157-8.
26. KIM, PHIL. MATLAB Deep Learning. *Springer*. 2017. Disponible en: DOI%2010 . 1007/978-1-4842-2845-6.
 27. BRESSEM, KENO K., ADAMS, LISA C., GAUDIN, ROBERT A., TRÖLTZSCH, DANIEL, HAMM, BERND, MAKOWSKI, MARCUS R., SCHÜLE, CHAN YONG, VAHLDIEK, JANIS L. AND NIEHUES, STEFAN M. Highly accurate classification of chest radiographic reports using a deep learning natural language model pre-trained on 3.8 million text reports. *Bioinformatics*. 2021, vol. 36, n.º 21, págs. 5255-5261. Disponible en: DOI%2010 . 1093/BIOINFORMATICS/BTAA668.
 28. JOSHI, RAMCHANDRA, GOEL, PURVI AND JOSHI, RAVIRAJ. DEEP LEARNING FOR HINDI TEXT CLASSIFICATION: A COMPARISON. Deep Learning for Hindi Text Classification: A Comparison. *Springer*. 2019, vol. 11886, n.º 94-101, págs. 5255-5261. Disponible en: DOI%2010 . 1007/978-3-030-44689-5_9.
 29. REYES, JULIO, RAMIREZ, DIEGO AND PACIELLO, JULIO. Automatic classification of source code archives by programming language: A deep learning approach. *International Conference on Computational Science and Computational Intelligence (CSCI)*. 2017, págs. 514-519. Disponible en: DOI%2010 . 1109/CSCI.2016.0103.
 30. PAOLETTI, ME, HAUT, JM, PLAZA, J AND PHOTOGRAMMETRY, A PLAZA. Deep learning classifiers for hyperspectral imaging: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*. [s.f.]. Disponible en: <https://doi.org/10.1016/j.isprsjprs.2019.09.006>.
 31. ZHAO, MINGHAO, HU, CHENGQUAN, WEI, FENGLIN, WANG, KAI, WANG, CHONG AND JIANG, YU. Real-Time Underwater Image Recognition with FPGA Embedded System for Convolutional Neural Network. *Sensors*. 2019, vol. 19, n.º 2, pág. 350. Disponible en: DOI % 2010 . 3390/S19020350.
 32. CHOW, JUN KANG, SU, ZHAOYU, WU, JIMMY, LI, ZHAOFENG, TAN, PIN SIANG, LIU, KUAN FU, MAO, XIN AND WANG, YU HSING. Artificial intelligence-empowered pipeline for image-based inspection of concrete structures. *Automation in Construction*. 2020, vol. 120, pág. 103372. Disponible en: DOI % 2010 . 1016/J. AUTCON. 2020. 103372.
 33. SHIDDIEQY, HASBI ASH, HARIADI, FARKHAD IHSAN AND ADIONO, TRIO. Implementation of deep-learning based image classification on single board computer. *International Symposium on Electronics and Smart Devices (ISESD)*. 2017, vol. 2018-January, págs. 133-137. Disponible en: DOI%2010 . 1109/ISESD.2017.8253319.
 34. LI, JIAHAO, LI, BIN, XU, JIZHENG, XIONG, RUIQIN AND GAO, WEN. Fully Connected Network-Based Intra Prediction for Image Coding. *IEEE Transactions on Image Processing*. 2018, vol. 27, n.º 7, págs. 3236-3247. Disponible en: DOI%2010 . 1109/TIP . 2018 . 2817044.
 35. DOGO, E. M., AFOLABI, O. J., NWULU, N. I., TWALA, B. AND AIGBAVBOA, C. O. A Comparative Analysis of Gradient Descent-Based Optimization Algorithms on Convolutional Neural Networks. *Proceedings of the International Conference on Computational Techniques, Electronics and Mechanical Systems*. 2018, págs. 92-99. Disponible en: DOI% 2010 . 1109/CTEMS . 2018 . 8769211.
 36. VALLEJO, WILLIAM, DÍAZ-URIBE, CARLOS AND FAJARDO, CATALINA. Google Colab and Virtual Simulations: Practical e-Learning Tools to Support the Teaching of Thermodynamics and to Introduce Coding to Students. *ACS Omega*. 2022, págs. 92-99. Disponible en: DOI % 2010 . 1021 / ACSOMEGA . 2C00362.
 37. RADIUK, PAVLO M. Impact of Training Set Batch Size on the Performance of Convolutional Neural Networks for Diverse Datasets. *Information Technology and Management Science*. 2018, vol. 20, n.º 1. Disponible en: DOI%2010 . 1515/ITMS-2017-0003.
 38. KESKAR, NITISH SHIRISH, NOCEDAL, JORGE, TANG, PING TAK PETER, MUDIGERE, DHEEVATSA AND SMELYANSKIY, MIKHAIL. On Large-Batch Training for Deep Learning: Generalization Gap and Sharp Minima. *5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings*. 2016. Disponible en: <https://arxiv.org/abs/1609.04836v2>.
 39. LI, WENMEI, LIU, HAIYAN, WANG, YU, LI, ZHUANGZHUANG, JIA, YAN AND GUI, GUAN. Deep Learning-Based Classification Methods for Remote Sensing Images in Urban Built-Up Areas. *IEEE Access*. 2016, vol. 7, págs. 36274-36284. Disponible en: DOI%2010 . 1109/ACCESS . 2019 . 2903127.

40. PANWAR, HARSH, GUPTA, P. K., SIDDIQUI, MOHAMMAD KHUBEB, MORALES-MENENDEZ, RUBEN AND SINGH, VAISHNAVI. Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet. *Chaos, Solitons Fractals*. 2020, vol. 138, pág. 109944. Disponible en: DOI%2010.1016/J.CHAOS.2020.109944.
41. MAXWELL, AARON E., WARNER, TIMOTHY A. AND GUILLÉN, LUIS ANDRÉS. Accuracy assessment in convolutional neural network-based deep learning remote sensing studies—part 1: Literature review. *Remote Sensing*. 2021, vol. 13, n.º 13. Disponible en: DOI%2010.3390/RS13132450.
42. TANER, ALPER, BENAL ÖZTEKIN, SIM, DURAN, HÜSEYİN, PAMPURO, NICCOLÒ, BIDDOCU, MARCELLA AND MANUEL MIRÁS-AVALOS, JOSÉ. Performance Analysis of Deep Learning CNN Models for Variety Classification in Hazelnut. *Sustainability*. 2021, vol. 13, n.º 12, pág. 6527. Disponible en: DOI%2010.3390/SU13126527.
43. VINAYAHALINGAM, SHANKEETH, KEMPERS, STEVEN, LIMON, LORENZO, DEIBEL, DIONNE, MAAL, THOMAS, HANISCH, MARCEL, BERGÉ, STEFAAN AND XI, TONG. Classification of caries in third molars on panoramic radiographs using deep learning. *Scientific Reports*. 2021, vol. 11, n.º 1, págs. 1-7. Disponible en: DOI%2010.1038/s41598-021-92121-2.
44. CRYSTAL, MARIA, OROZCO, E AND REBONG, CORAZON B. Vehicular Detection and Classification for Intelligent Transportation System: A Deep Learning Approach Using Faster R-CNN Model. 2021. Disponible en: DOI%2010.5013/IJSSST.a.20.S2.11.
45. DANDAVATE, RUCHA AND PATODKAR, VINEET. CNN and data augmentation based fruit classification model. *Proceedings of the 4th International Conference on IoT in Social, Mobile, Analytics and Cloud, ISMAC 2020*. 2020, págs. 784-787. Disponible en: DOI%2010.1109/I-SMAC49090.2020.9243440.
46. LINGWAL, SURABHI, BHATIA, KOMAL KUMAR AND TOMER, MANJEET SINGH. Image-based wheat grain classification using convolutional neural network. *Multimedia Tools and Applications*. 2021, vol. 80, n.º 28–29, págs. 35441-35465. Disponible en: DOI%2010.1007/S11042-020-10174-3/METRICS.
47. HUANG, NEN FU, CHOU, DONG LIN, LEE, CHIA AN, WU, FENG PING, CHUANG, AN CHI, CHEN, YI HSIEN AND TSAI, YIN CHUN. Smart agriculture: real-time classification of green coffee beans by using a convolutional neural network. *IET Smart Cities*. 2020, vol. 2, n.º 4, págs. 167-172. Disponible en: DOI%2010.1049/IET-SMC.2020.0068.



Artículo de **libre acceso** bajo los términos de una **Licencia Creative Commons Reconocimiento – NoComercial – CompartirIgual 4.0 Internacional**. Se permite que otros remezcLen, adapten y construyan a partir de su obra sin fines comerciales, siempre y cuando se otorgue la oportuna autoría y además licencien sus nuevas creaciones bajo los mismos términos.