

Estudio Analítico de desempeño entre Redes Neuronales tipo Perceptrón y Convolucionales para el reconocimiento de imágenes

An analytical study of the performance of Perceptron type and Convolutional Neural Networks for image recognition

Luis Arturo Medina-Muñoz ^{a,*}, Luis Carlos Valenzuela-Soqui ^b, Samuel Gonzalez-Lopez ^a, Felipe Morales-Solis ^c, Rocio Alejandrina Valenzuela-Salomon ^c

^a Departamento de Posgrado e Investigación, Instituto Tecnológico de Nogales, Avenida Instituto Tecnológico # 911, Col. Granjas. C.P. 84065, Nogales, Sonora, México.

^b Departamento de Sistemas y Computación, Tecnológico Nacional de México, Instituto Tecnológico de Nogales, Avenida Instituto Tecnológico # 911, Col. Granjas. C.P. 84065, Nogales, Sonora, México.

^c Departamento de Ciencias Económico Administrativa, Tecnológico Nacional de México, Instituto Tecnológico de Nogales, Avenida Instituto Tecnológico # 911, Col. Granjas. C.P. 84065, Nogales, Sonora, México.

Correo electrónico: luis.mm@nogales.tecnm.mx

(Recibido: 20 de marzo 2025; Aceptado: 28 de abril 2025; Publicado: 01 de mayo 2025)

Resumen

El presente trabajo muestra el comportamiento analítico de 2 estructuras distintas de Redes Neuronales (RN) para el reconocimiento de imágenes de verduras como el chile, tomate y cebolla, implementado en un sistema embebido Raspberry Pi®. El objetivo que se persigue con este estudio es definir un algoritmo general que pueda reconocer imágenes sin necesidad de ajustar parámetros, implementado en un sistema embebido. Este algoritmo puede ser seleccionado de acuerdo con los resultados que arroje este estudio. También, ofrecer a la industria de procesamiento de la fruta y verdura, instalados en el área fronteriza, un dispositivo barato y confiable de reconocimiento para automatizar sus procesos de selección y empaque, así como de extender el uso de estos resultados a la industria manufacturera regional para automatizar el proceso de inspección de calidad.

Palabras claves: Algoritmo, Imágenes, Industria, Parámetros, Sistemas, Visión.

Abstract

This paper presents the analytical behavior of two different Neural Network (NN) structures for image recognition of vegetables such as chili peppers, tomatoes, and onions, implemented on a Raspberry Pi® embedded system. The objective of this study is to define a general algorithm capable of recognizing images without requiring adjustments to parameters implemented in embedded systems. This algorithm can be selected based on the results of this study. Furthermore, the aim is to offer the fruit and vegetable processing industry in the border region an inexpensive and reliable recognition device to automate their selection and packaging processes, as well as to extend the application of these results to the internal processes of the regional manufacturing industry to automate quality inspection.

Keywords: Algorithm, Images, Industry, Parameters, Systems, Vision.

1. Introducción

La inteligencia artificial (IA) ofrece muchas oportunidades para contribuir al bienestar de las personas y al progreso de las economías y las sociedades, pero también una variedad de nuevos desafíos éticos, legales, sociales y tecnológicos. [1].

Con el constante desarrollo y aplicación de algoritmos de inteligencia artificial (IA) han aparecido alternativas de bajo costo para su procesamiento rápido y confiable. Los mejores sistemas de procesamiento embebido son caros por lo que están fuera del alcance de las instituciones educativas que realizan este tipo de investigaciones.

La clasificación de imágenes es un problema fundamental en el campo de la visión artificial. Con el rápido desarrollo de internet y la inteligencia artificial, se genera una gran cantidad de datos de imagen a diario [2].

Raspberry Pi® es un claro ejemplo de esto, pero también se ha desarrollado software capaz de acoplarse a estos sistemas embebidos y poder realizar tareas con herramientas de inteligencia artificial. *Software* como Google colab®, en lenguaje de programación de Python®, es un ambiente de programación del cual un programador puede hacer uso para desarrollar aplicaciones inteligentes en sistemas embebidos, resultando esto en un ahorro de tiempo y recursos.

2. Desarrollo

Los Sistemas Inteligentes en la actualidad son cada vez más utilizados tanto en la vida cotidiana como en las actividades de agricultura. El uso de las arquitecturas de Redes Neuronales Convolucionales (CNN) mediante la transferencia de aprendizaje permite el uso de modelos de clasificación más robustos y facilitan enormemente la etapa de sintonización de hiperparámetros y entrenamiento de nuevas clases [3].

En [4] se desarrollan técnicas de Deep Learning para la implementación de una Red Neuronal Convolucional para el reconocimiento de acciones humanas. La aplicación de esta red se puede trasladar a sistemas de vigilancia remotos o embebidos para realizar el monitoreo de comportamientos, acciones agresivas, prevención de delitos entre otras más.

La Red Neuronal Convolucional (CNN) es la más utilizada entre 9 tipos de Redes Neuronales Artificiales (RNA), y en el procesamiento de imágenes se diferencian dos fases primordiales, pre-procesamiento y entrenamiento. En conclusión, la CNN tiene los mejores resultados por sus altos porcentajes de precisión y su implementación es cada vez mayor en el campo de la visión artificial [5].

La clasificación de frutas mediante la Inteligencia Artificial ha sido un factor muy importante en la optimización de procesos industriales en el sector alimentario.

Esta investigación es importante porque permite enmarcar el punto inicial para el desarrollo de sistemas autónomos usando visión artificial con pocos recursos computacionales, cuyo objetivo es entrenar un modelo de visión artificial utilizando una Red Neuronal Convolucional para la clasificación de frutas [6].

En [7] se describe una técnica de aprendizaje profundo para clasificar imágenes reales y generadas por IA basada en la arquitectura MobileNetV2. Los experimentos demuestran la capacidad del modelo para reconocer imágenes generadas por IA con alta precisión y un rendimiento de clasificación eficiente.

Los objetivos principales de esta investigación son 2 principalmente, el primero, que los resultados que se obtengan nos lleven a definir un algoritmo que sea capaz de poder aplicarse a cualquier conjunto de imágenes para su reconocimiento. Segundo, que la industria de manufactura vea como las instituciones de educación superior son capaces de proveer desarrollos tecnológicos baratos y confiables para sus procesos.

El presente trabajo muestra un estudio analítico entre 2 diferentes estructuras neuronales para reconocimiento de imágenes implementados en un sistema embebido.

3. Metodología

Este trabajo muestra un procedimiento de implementación de algoritmos de inteligencia artificial en un sistema embebido Raspberry Pi® para entrenar 2 tipos de redes neuronales, el Perceptrón tradicional y lo que es una red neuronal del tipo convolucional.

Los resultados servirán para hacer un análisis y determinar cuál aplicación es la más eficiente para este tipo de reconocimiento de imágenes. Para implementar

esta propuesta se siguió un flujo de trabajo como el que se muestra en la figura 1.

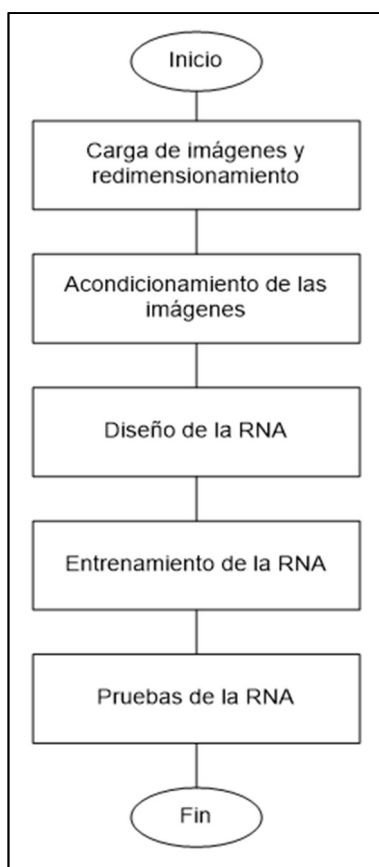


Fig. 1 Flujo de trabajo de la propuesta.

Para la realización de este estudio analítico se tomaron alrededor de 150 imágenes con una cámara convencional de 3 productos distintos, chile, tomate y cebolla. Aunque es posible incrementar la cantidad de imágenes de entrenamiento hay que estar conscientes que entre mayor sea la cantidad de imágenes más tiempo tarda el sistema para entrenar.

3.1 Perceptrón propuesto

La estructura del Perceptrón propuesto lo podemos observar en la figura 2, consta de 3 capas:

- Capa de entrada con n neuronas determinadas por el tamaño de la imagen.
- 256 y 128 neuronas en una primera y segunda capa oculta, estas cantidades se determinaron a través de una constante experimentación, entrenando en diversas ocasiones a la red neuronal.
- Una capa de salida con 3 neuronas, donde cada una corresponde al parámetro identificado.

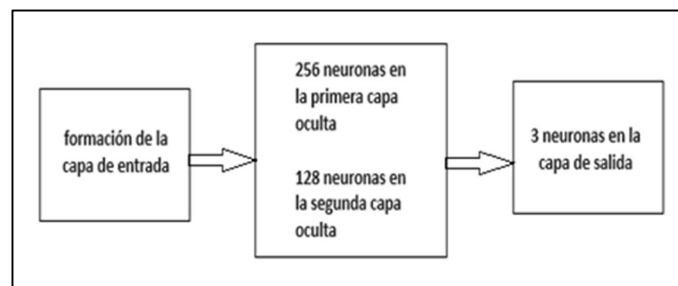


Fig. 2 Estructura del Perceptrón propuesto.

Para la creación de la capa de entrada, se redujo el tamaño de las imágenes a (200, 200, 3) respetando los 3 colores base RGB indicados por el valor del último argumento.

Las imágenes como tal no pueden definir una cantidad de neuronas determinadas para esta capa, por lo que utiliza un método contenido en la librería de keras, llamado FLATTEN, el cual con la información del tamaño de la imagen forma automáticamente la cantidad de neuronas necesarias en la primera capa mediante la conformación de un arreglo binario unidimensional que contiene la información de la imagen.

Luego, mediante un proceso de experimentación constante de entrenamiento de la red neuronal y evaluación de resultados se determina el uso de 2 capas ocultas, una primera de 256 neuronas y una segunda de 128 con funciones de activación lineal.

Finalmente, como son 3 diferentes imágenes a ser reconocidas, la capa de salida consta de 3 neuronas, una que determina el porcentaje de reconocimiento para imágenes de chile, una segunda para determinar el grado de reconocimiento del tomate y finalmente la última determinando el valor con la que la cebolla es reconocida.

3.2 Red neuronal convolucional

La figura 3 muestra una estructura del modelo creado para una red neuronal convolucional, al igual que el perceptrón la determinación de la cantidad de neuronas en una capa intermedia se obtiene a través de la experimentación y análisis de datos.

La cantidad de neuronas en la capa inicial es determinada al igual que en el modelo de perceptrón por el método FLATTEN y la cantidad de neuronas en la capa de salida corresponden a la cantidad de diversas imágenes a ser reconocidas.

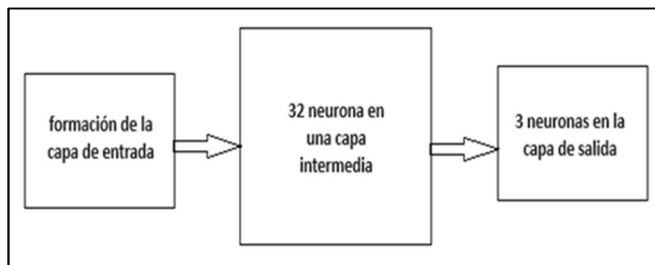


Fig. 3 Estructura de red neuronal convolucional.

Ambos modelos se compilaron utilizando el mismo optimizador, ADAM®, y una cantidad de épocas iguales, 100.

4. Resultados y discusión

El sistema entrenado se probó con el equivalente al 10% del total de las imágenes utilizadas en el entrenamiento, las cuales fueron 150 en total, los resultados se muestran en la tabla 1.

Tabla 1 Porcentaje de reconocimiento de verduras por RNA tipo Perceptrón.

Clase	Porcentaje de reconocimiento
Tomate	97.06 %
Chile	94.88 %
Cebolla	97.94 %

La tabla 1 muestra los resultados obtenidos en términos de porcentajes, de acuerdo con los valores mostrados, la RNA indica que hay un error de reconocimiento de aproximadamente 5 % en el chile, en el tomate y cebolla muestra un error de reconocimiento de aproximadamente 3 %. Si consideramos un valor promedio de reconocimiento, vemos que el 96.62 % de las veces el sistema trabaja de forma adecuada.

La tabla 2 muestra la Matriz de confusión propuesta, en esta se utilizaron las siguientes abreviaciones:

- **VP:** Es una verdura correcta dentro del conjunto de pruebas y el sistema la reconoce de manera adecuada
- **VN:** El sistema confunde, por ejemplo, una verdura con otra
- **FP:** Se presenta al sistema una verdura real y no la identifica correctamente
- **FN:** No es una verdura real y el sistema no la reconoce como tal.

Tabla 2 Matriz de confusión propuesta.

	Predicción	
Real	0	1
0	FN	VN
1	FP	VP

La tabla 3 muestra los datos del proceso de validación cruzada de la Red Neuronal, utilizando 15 imágenes para establecer el valor de exhaustividad que no dará una idea de que tan eficiente es para el reconocimiento de imágenes. Estos datos se distribuyeron de acuerdo con la propuesta indicada en la tabla 2.

Tabla 3 Matriz de confusión para RNA Perceptrón.

	Predicción	
Real	0	1
0	0	2
1	0	13

En la tabla 3 los resultados que se muestran indican que, de las 15 imágenes de verduras, tomate, chile y cebolla, 13 fueron reconocidas de manera correcta y hubo 2 verduras, que son los verdaderos negativos, en donde el sistema las confundió entre sí, es decir, un tomate lo confundió con una cebolla, etc.

Con los datos obtenidos en la tabla 3 se calcula la eficiencia de la RNA, que se muestra mediante el cálculo de la exhaustividad:

$$Exhaustividad = \frac{13}{13 + 2} = 0.86$$

La RNA tipo Perceptrón propuesta, muestra que tiene un 86 % de eficiencia para reconocer las 3 diferentes clases de verduras usadas para este trabajo de análisis.

La tabla 4 muestra los resultados obtenidos en términos de porcentajes para la RNA del tipo convolucional, de acuerdo con los valores mostrados, la RNA indica que hay un error de reconocimiento de aproximadamente 3 %, el cual es casi uniforme para cada una de las verduras clasificadas. Si consideramos un valor promedio de reconocimiento, vemos que el

97.56 % de las veces el sistema clasifica de forma adecuada a cada una de las verduras usadas en este proceso de clasificación.

Tabla 4 Porcentaje de reconocimiento de RNA tipo convolucional.

Clase	Porcentaje de reconocimiento
Chile	97.16 %
Tomate	97.69 %
Cebolla	97.84 %

La tabla 5 muestra como de los 15 datos usados como parte del proceso de validación cruzada, 14 fueron identificados de manera correcta, hubo un dato identificado de forma incorrecta, es decir, un falso positivo. De igual manera, no hay datos catalogados como falsos negativos y verdaderos negativos.

Tabla 5 Matriz de confusión para RNA convolucional.

	Predicción	
Real	0	1
0	0	1
1	0	14

Con los datos obtenidos en la tabla 5 se calcula la eficiencia de la RNA, que se muestra mediante el cálculo de la exhaustividad

$$\text{Exhaustividad} = \frac{14}{14 + 1} = 0.93$$

La RNA del tipo convolucional propuesta, muestra que tiene un 93 % de eficiencia para reconocer las 3 diferentes clases de verduras usadas para este trabajo de análisis.

5. Conclusiones

La utilización de sistemas embebidos, como Raspberry Pi®, para diferentes tareas de procesamiento de información todavía se usa con cierto cuidado cuando se trata de implementación de procesos de IA como las Redes Neuronales en conjunto con el uso de sistemas de visión, debido a sus pocos recursos de procesamiento. Esto ha venido cambiando con el uso de librerías de procesamiento de IA como keras, tensor Flow® y

algunos algoritmos de RNA convolucionales que han permitido que estos sistemas embebidos puedan usarse para ciertas tareas de reconocimiento artificial.

La tabla 1 muestra los resultados de la implementación de una RNA tipo Perceptrón con una capa de entrada, una oculta y una de salida, con un porcentaje promedio de reconocimiento de 96.62 %.

La tabla 4 muestra los resultados de la implementación de una RNA tipo convolucional, con un porcentaje promedio de reconocimiento de 97.56 %. La eficiencia mostrada por cada una de las implementaciones de RNA indican que la red tipo convolucional tiene un mayor grado de eficiencia con un 93 % comparado con el 86 % de la red tipo Perceptrón.

El análisis numérico de estas implementaciones muestra que existen diferencias en cuanto al reconocimiento de los patrones utilizados en este estudio, no se puede afirmar si el comportamiento analítico seguirá siendo el mismo para otro tipo de patrones, ya que es necesario ajustar los parámetros de los algoritmos utilizados.

De manera general, el presente trabajo muestra que es posible utilizar algoritmos de IA, herramientas de visión en sistemas embebidos con pocas capacidades de procesamiento en el reconocimiento de imágenes y obtener resultados muy buenos con alrededor de un 3 % promedio de error.

Trabajos futuros implican el desarrollar un algoritmo general que sea capaz de identificar imágenes con el solo hecho de cargar los datos de entrenamiento necesarios si necesidad de realizar ajustes en los diferentes procesos y estructuras de las RNA propuestas.

6. Agradecimientos

Agradecemos al Tecnológico Nacional de México campus Nogales por todas las facilidades prestadas para el desarrollo de este trabajo.

7. Referencias

- [1] Lins, S., Pandl, K.D., Teigeler, H. Artificial Intelligence as a Service. Bus Inf Syst Eng 63, 441–456 (2021). <https://doi.org/10.1007/s12599-021-00708-w>.

[2] Jiang li Liu, Jin feng Li, Guang yan Huang. Research on Deep Learning-based Image Processing and Classification Techniques for Complex Networks. Applied Mathematics and Nonlinear Sciences, 10(1) (2025) 1-22, China.

[3] Deny Lizbeth Hernández Rabadat, Juan Paulo Sánchez Hernández, Sandra Elizabeth León Sosa, Juan Frausto Solís y Juan Javier González Barbosa. “Análisis de arquitecturas de redes neuronales convolucionales aplicadas en la detección de cenicilla polvorienta en cultivos de jitomate”. Departamento de Tecnologías de Información, Universidad Politécnica del estado de Morelos. VII Congreso Nacional de Riego, Drenaje y Biosistemas. Noviembre del 2022.

[4] Mitchell Ángel Gómez Ortega. “Implementación de una red neuronal convolucional para el reconocimiento de acciones humanas”. Universidad Politécnica del Valle De México, División de Ingeniería Mecatrónica y Mecánica Automotriz. Research in Computing Science 150(6), 2021.

[5] Jorge Saa-Saltos, Llanos Mora-López, José del Campo-Ávila. “Análisis de redes neuronales artificiales para el reconocimiento automático de plantas”. Universidad Técnica Estatal de Quevedo, Ecuador. Revista Ibérica de Sistemas y Tecnologías de Información. RISTI, No. E52, 08/2022.

[6] Jonathan Víctor Aguilar-Alvarado, Milton Alfredo Campoverde-Molina. “Clasificación de frutas basadas en redes neuronales convolucionales”. Unidad Académica de Tecnologías de la Información y la Comunicación (TIC), Jefatura de Posgrados Universidad Católica de Cuenca, Cuenca, Ecuador. Polo del conocimiento. Edición núm. 41, Vol. 5, No 01, enero 2020, pp. 3-22, ISSN: 2550 - 682X.

[7] Jaffrin, L. C.; Anvesh, K.; Balaji Venkata Ratnam, Y.; Nikhileshwar Reddy, V. (2025) “AI Vs Real Image Classification Using Deep Learning”. Proceedings of the 4th International Conference on Information Technology, Civil Innovation, Science and Management (ICITSM 2025). DOI: 10.4108/eai.28-4-2025.2358080.