



UNIVERSIDAD CÉSAR VALLEJO

**FACULTAD DE INGENIERÍA Y ARQUITECTURA
ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS**

**Redes Neuronales Convolucionales para determinar la acidez y
grados Brix de Pitahaya Amarilla (*Hylocereus megalanthus*) en la
Región Amazonas**

TESIS PARA OBTENER EL TÍTULO PROFESIONAL DE:

Ingeniero de Sistemas

AUTOR:

Cueva Caro, Jeison Enrique (orcid.org/0000-0001-8482-3407)

ASESOR:

Dr. Necochea Chamorro, Jorge Isaac (orcid.org/0000-0002-3290-8975)

LÍNEA DE INVESTIGACIÓN:

Sistemas de Información y Comunicaciones

LÍNEA DE RESPONSABILIDAD SOCIAL UNIVERSITARIA:

Desarrollo económico, empleo y emprendimiento

LIMA - PERÚ

2022

DEDICATORIA

En primer lugar, a Jehová Dios por
brindarnos salud, trabajo y un día mas
de vida en estos tiempos difíciles de pandemia

A mi Madre que no pierde las esperanzas de
ver a sus hijos para formarse como personas
y profesionales sobre todo por su amor infinito.

También, a la memoria de mi padre,
A mi hijo, a su mamá y mis hermanas
por su motivación y apoyo durante
todo el aprendizaje de la carrera.

AGRADECIMIENTO

Definitivamente agradezco a mi familia en general por brindarme siempre los ánimos y no dejarme rendir de igual manera agradezco a los docentes y mis asesores, al Dr. Jorge Necochea Chamorro por su apoyo, conocimiento y colaboración para culminar el desarrollo de esta investigación.

un agradecimiento especial al PhD. Ives Yóplac Tafur por su asesoría que brindó para permitir el desarrollo de esta investigación, así mismo al Ing. Julio Cesar Santos Peláez, por ayudar a realizar las evaluaciones de pitahaya.

A la Universidad Nacional Toribio Rodríguez de Mendoza de Amazonas - UNAT-A por permitir el uso de sus laboratorios de Bromatología y así cumplir con los objetivos propuestos.

ÍNDICE DE CONTENIDOS

DEDICATORIA	ii
AGRADECIMIENTO	iii
RESUMEN	ix
ABSTRACT	x
I. INTRODUCCIÓN	1
II. MARCO TEÓRICO	6
III. METODOLOGÍA	21
3.1. Tipo y diseño de investigación.....	21
3.1.1. Tipo de investigación	21
3.1.2. Diseño de investigación	21
3.2. Variables y operacionalización	21
3.3. Población, muestra y muestreo	22
3.3.1. Población	22
3.3.2. Unidad de análisis.....	23
3.4. Técnicas e instrumentos de recolección de datos	23
3.5. Procedimientos.....	23
3.5.1. Procedimientos en la obtención de imágenes y pruebas de laboratorios.....	25
3.5.2. Limitaciones	35
3.6. Métodos de análisis de datos	35
3.7. Aspectos éticos.....	35
IV. RESULTADOS	37
4.1. Resultados Descriptivos de la Investigación.....	37
4.1.1. Análisis de la Acidez Titulable, Grados Brix y pH.....	37
4.1.2. Configuración y entrenamiento de la red neuronal.....	43

4.1.3. Prueba de la red, exactitud y tiempo promedio de respuesta	44
4.2. Resultados Inferenciales de la Investigación.....	47
V. DISCUSIÓN	55
VI. CONCLUSIONES	57
VII. RECOMENDACIONES	59
REFERENCIAS	60
ANEXOS	64

ÍNDICE DE TABLAS

Tabla 1. Métricas de evaluación de resultados	15
Tabla 2. Tabla de frecuencia Acidez Titulable	32
Tabla 3. Tabla de frecuencia Grados Brix	33
Tabla 4. Tabla de frecuencia pH	34
Tabla 5. Resultados de Hipótesis	54

ÍNDICE DE FIGURAS

Figura 1. Ejemplo de operación de convolución	13
Figura 2. Aplicación de la función de activación	13
Figura 3. Ejemplo de capa de agrupamiento	14
Figura 4. Ejemplo de Pitahaya Amarilla	16
Figura 5. Etapas de la maduración de la Pitahaya Amarilla	18
Figura 6. Evolución de los cambios físicos en la Pitahaya Amarilla	19
Figura 7. Evolución de los cambios químicos en la Pitahaya Amarilla	20
Figura 8. Diagrama de desarrollo del proyecto	24
Figura 9. Clasificación de pitahaya en niveles de maduración 0 al 6	26
Figura 10. Clasificación de pitahaya en niveles de maduración 0 al 6	26
Figura 11. Procedimiento para evaluación de Grados Brix	27
Figura 12. Evaluación del pH	28
Figura 13. Evaluación de la acidez titulable	29
Figura 14. Escala de color elaborada según maduración de la fruta	29
Figura 15. Importación de librerías y preprocesamiento de la data	30
Figura 16. Evaluación de Acidez Titulable	33
Figura 17. Evaluación de grados Brix	34
Figura 18. Evaluación del pH	35
Figura 19. Promedios de Grados Brix ($^{\circ}\text{Bx}$) – Acidez Titulable (AT) – pH	36
Figura 20. Fórmula para determinar el índice de madurez	37
Figura 21. Índice de madurez de pitahaya	37
Figura 22. Épocas de entrenamiento de la red	38
Figura 23. Resultados del entrenamiento de la red	39
Figura 24. Configuración fase de prueba de la red	39
Figura 24. Impresión de resultados por la red neuronal	40
Figura 26. Tiempo empleado en la predicción	40
Figura 27. Resultados y matriz de confusión obtenida	41
Figura 28. Exactitud de las predicciones	41
Figura 29. Prueba de bondad Kolmogorov – Acidez Titulable	43
Figura 30. Prueba de bondad Kolmogorov – Grados Brix	43
Figura 31. Prueba de bondad Kolmogorov – pH	43

Figura 32. Análisis de Varianza de Acidez Titulable	44
Figura 33. Gráfico de barras – Acidez Titulable	44
Figura 34. Análisis de Varianza de Grados Brix	45
Figura 34. Gráfico de barras – Grados Brix	46
Figura 34. Análisis de Varianza de PH	47
Figura 34. Gráfico de barras – PH	47
Figura 34. Requisitos químicos de la pitahaya amarilla	48

RESUMEN

En la presente investigación se desarrolló una red neuronal convolucional con el objetivo de determinar la Acidez y Grados Brix de la Pitahaya Amarilla (*Hylocereus megalanthus*), de manera que se facilite la identificación de la fruta según su estado de madurez. Metodológicamente, la investigación fue de tipo aplicada y diseño experimental, empleando un enfoque cuantitativo para comparar el nivel de exactitud y tiempo de respuesta logrado con la aplicación de la mencionada red. Para ello, la muestra estuvo conformada por 12 pitahayas por cada nivel de madurez (nivel 0 a nivel 6), de las cuales se tomaron un total de 1620 fotografías para conformar la base de datos, utilizando en 80% de estas para el entrenamiento y dejando el 20% restante para la validación de resultados. Simultáneamente, se realizaron pruebas de laboratorio para determinar la acidez (AT), grados Brix (SST) y pH, empleando un refractómetro Ph Meter y Bureta Digital, lo que permitió caracterizar los valores de tales indicadores para cada nivel de maduración. En cuanto a la puesta a prueba de la red, se obtuvo una exactitud del 100% en toda la evaluación, la cual demoró un promedio de 0.8 segundos por imagen analizada. Así, la red resulta eficaz y eficiente para predecir el nivel de madurez de la Pitahaya Amarilla.

Palabras clave: Red neuronal convolucional, pitahaya, refractómetro, grados Brix, acidez titulable.

ABSTRACT

In the present investigation, a convolutional neural network was developed with the objective of determining the Acidity and Brix Degrees of the Yellow Pitahaya (*Hylocereus megalanthus*), in order to facilitate the identification of the fruit according to its state of maturity. Methodologically, the research was of an applied type and experimental design, using a quantitative approach to compare the level of accuracy and the response time achieved with the application of the aforementioned network. For this, the sample consisted of 12 pitahayas for each level of maturity (level 0 to level 6), of which a total of 1640 photographs were taken to form the database, using 80% of these for training and leaving the remaining 20% for validation of results. Simultaneously, laboratory tests were carried out to determine the titratable acidity (AT), Brix degrees (SST) and pH, using a refractometer, which allowed characterizing the values of such indicators for each level of maturation. Regarding the testing of the network, an accuracy of 100% was obtained in the entire evaluation, which demonstrated an average of 0.8 seconds per image analyzed. Thus, the network is effective and efficient to predict the level of maturity of the Yellow Pitahaya.

Keywords: Convolutional neural network, pitahaya, refractometer, degrees Brix, titratable acidity.

I. INTRODUCCIÓN

Los avances tecnológicos han transformado la dinámica de la sociedad desde la antigüedad, pero ha sido en los últimos años que este efecto se ha visto incrementado en proporciones más allá de lo previsto, encabezado por el uso cotidiano de ordenadores para las tareas de cómputo y teniendo como máximos exponentes los últimos adelantos en inteligencia artificial, la robótica y la automatización (1). En este sentido, las tareas que en principio requerían una gran capacidad de cálculo y tiempo para un ser humano han sido delegadas a sistemas computarizados que ejecutan una serie de algoritmos para hallar la respuesta con un desempeño miles de veces superior (2).

De las técnicas anteriormente mencionadas, la inteligencia artificial ha jugado un papel fundamental en la optimización de procesos industriales, aumentando la eficiencia y eficacia a la vez que se reducen los errores humanos; bajo este panorama, destaca la aplicación de redes neuronales convolucionales para el análisis, clasificación y segmentación de imágenes (3, 4). El rango de aplicaciones de dicha tecnología ha sido sumamente amplio, desde la identificación de lugares geográficos (5), imágenes procedentes de exámenes médicos (6) o la identificación y clasificación de especies animales o vegetales (7, 8, 9), entre otros fines. Este último campo de aplicación resulta especialmente útil para la actividad agrícola, en especial, lo relativo a clasificación de productos según estándares de calidad (10, 11).

Entre tanto, la Pitahaya Amarilla es considerada como una fruta exótica, de sabor entre dulce y agridulce, originaria de América Central y América del Sur, que goza de una gran demanda en el mercado nacional e internacional debido a excelentes cualidades nutricionales y sensoriales, destacando un alto contenido de ácido ascórbico (vitamina C) y antioxidantes, tiene un efecto antiinflamatorio y favorece al alivio del estrés, siendo incluida en dietas para tratar problemas estomacales, para contrarrestar la obesidad, la diabetes e incluso algunos tipos de cáncer; tales propiedades le confieren un costo de venta alto en estado fresco, de lo que se desprende una alta rentabilidad de producción, siendo cultivada comercialmente en México, Ecuador, Colombia, Nicaragua, Guatemala, Estados Unidos, Israel y también en el Perú, especialmente después del 2010, cuando

empezó a ganar popularidad en el ambiente gastronómico y comercial al conocerse en mayor medida sus bondadosas propiedades (11, 12, 13).

Aunado a lo anterior, (14) establecen que el comportamiento del mercado se muestra favorable al consumo de pitahaya en distintos países del mundo, especialmente América del norte y Europa, por lo que la producción de esta fruta se ha visto incrementada en los últimos años para satisfacer la demanda de los consumidores, dónde, se ha detectado la importancia que tiene la consistencia de la apariencia y el sabor; de allí que varios estudios se enfoque en proponer métodos de selección basados en los índices de maduración de la fruta.

Es así, que, debido a su potencial económico, algunos países ya han enfocado esfuerzos para normalizar su producción, como el caso de Colombia, quien desde el año 1996 posee una norma técnica emanada del instituto colombiano de normalización (COINTEC), dónde expone las pautas técnicas que regulan el uso de la Pitahaya Amarilla como materia prima, recomendaciones técnicas procesamiento y requisitos de exportación (15).

En el contexto nacional, la problemática actual reside en el desconocimiento de los productores locales de Pitahaya Amarilla sobre la manera de determinar el momento oportuno para cosechar la fruta. Si bien, en un nivel técnico, la madurez puede ser determinada a partir de indicadores como la acidez (4.3 a 4.7) y los grados Brix (12° a 14°), su determinación en campo resulta en una tarea apoteósica para los productores locales, quienes no tienen las herramientas ni la formación técnica apropiada para conducir los ensayos de laboratorio requeridos, a lo cual se ha de sumar los costos que se incurriría al proceder de tal manera (13). Tradicionalmente, las técnicas de cosecha artesanal se basan en un cúmulo de experiencias previas de los productores, quienes por medio del ensayo y error han aprendido las características que deben presentar los cultivos para determinar el momento oportuno de la recolección, tales como la coloración y el tamaño de la fruta (11). Sin embargo, en el contexto de la presente investigación, se tiene que la Pitahaya Amarilla es una especie relativamente nueva en la región de Amazonas, por lo que los productores locales carecen de las bases necesarias para distinguir los signos que indican que la fruta está en el punto óptimo para ser colectada, lo que repercute en la efectividad de la cosecha y en la calidad de la fruta comercializada.

Es por ello, que el enfoque de la presente investigación es el de elaborar una herramienta que facilite dicha tarea, con un nivel de sencillez que le permita ser usada fácilmente por los productores locales y cuya efectividad favorezca el incremento de los niveles de producción y la calidad de la fruta cosechada. De esta manera, se propone desarrollar un modelo en el procesamiento de imágenes que a su vez emplea redes neuronales convolucionales, con la cual, en siguientes investigaciones se pueda seguir implementando una aplicación móvil o web para que los productores de la Pitahaya Amarilla puedan tomar una fotografía que pueda ser procesada para obtener una estimación certera sobre el estado de madurez de la fruta y así determinar si se encuentra en el momento óptimo para realizar su cosecha. Entre las principales ventajas de la propuesta se espera tener un método de análisis práctico de baja complejidad de operación para los productores agrícolas, la inmediatez de los resultados y un costo de implementación menor que el requerido para los exámenes de laboratorio tradicionales.

Por las razones antes expuestas, se presenta el **problema principal** de la investigación: ¿De qué manera se puede utilizar una red neuronal convolucional para determinar la Acidez (% ácido cítrico) y Grados Brix (cantidad °Bx - SST) de la Pitahaya Amarilla en la región Amazonas? A partir de ello, se derivan como **problemas específicos**: a) ¿Qué parámetros de SST, AT y pH presenta la Pitahaya Amarilla según los estados de maduración? b) ¿Qué nivel de exactitud se obtiene al utilizar una red neuronal convolucional para determinar la Acidez (% ácido cítrico) y Grados Brix (cantidad °Bx - SST) de la Pitahaya Amarilla en la región Amazonas?, y c) ¿Cuál es el tiempo de respuesta promedio al utilizar una red neuronal convolucional para determinar la Acidez y Grados Brix de la Pitahaya Amarilla en la región Amazonas?

Considerando lo anterior, la investigación presenta una **justificación práctica**, orientando una solución tecnológica a un problema palpable que en estos momentos afecta a la comunidad de productores de la Pitahaya Amarilla al limitar la productividad de la cosecha y la calidad general de los productos comercializados. En este sentido, la aplicación busca facilitar el trabajo de identificar el estado de la fruta para determinar si esta se encuentra o no en el momento oportuno para ser cosechada, maximizando la productividad obtenida. Se debe tomar en cuenta que el grado de maduración repercute en el desarrollo de

sus cualidades nutricionales y diversas propiedades como el tamaño final de la fruta, porcentaje de pulpa comestible, sabor, aroma, firmeza, entre otros, lo que influye en su valor comercial individual, e indirectamente, en la rentabilidad de la actividad para el productor.

Como **justificación social**, la investigación tiene el potencial de ayudar a la comunidad de productores de la Pitahaya Amarilla en la región de Amazonas a llevar a cabo sus labores de manera eficaz y eficiente, tomando en consideración que para muchos este es su principal ingreso económico. De manera similar, el desarrollo de la aplicación se justifica por cuanto provee un medio alternativo a los métodos de ensayo tradicional por medio de análisis de laboratorios, cuyo costo dificulta su aplicación de manera práctica y a gran escala.

En cuanto a la **justificación tecnológica**, el desarrollo de la investigación pone en manifiesto el uso de las últimas técnicas y adelantos de la informática para el desarrollo de soluciones a problemas reales que aquejan a la comunidad general, teniendo el caso particular de ahondar en la tecnología de las redes neuronales convolucionales. Los resultados a obtener no solo ilustrarán el éxito de la tecnología, sino que servirá de base como referencia a otros investigadores en contextos similares. Además, las técnicas y procedimientos a aplicar pueden ser de utilidad no solo para el caso del tema de estudio, centrado en la Pitahaya Amarilla, sino para otros tipos de frutos o cultivos, y en última instancia, a la agricultura en general.

De esta forma, la investigación persigue como **objetivo general**: Evaluar el uso de Redes Neuronales Convolucionales para determinar la acidez y grados Brix de Pitahaya Amarilla en la Región Amazonas. En función de ello, se tiene como objetivos específicos: a) Determinar los parámetros SST, Acidez Titulable y pH de la Pitahaya Amarilla según los estados de maduración, b) Determinar el nivel de exactitud al utilizar una red neuronal convolucional para determinar la Acidez y Grados Brix de la Pitahaya Amarilla en la región Amazonas, y c) Determinar el tiempo de respuesta promedio al utilizar una red neuronal convolucional para determinar la Acidez y Grados Brix de la Pitahaya Amarilla en la región Amazonas.

Para la presente investigación se formula la siguiente **hipótesis general**: Las Redes Neuronales Convolucionales permiten determinar eficientemente la acidez y grados Brix de Pitahaya Amarilla en la Región Amazonas, teniendo como

hipótesis específicas: a) Es posible determinar el estado de maduración de la fruta según los parámetros de, SST, Acidez Titulable y pH, b) El nivel de exactitud al utilizar una red neuronal convolucional para determinar los Grados Brix y Acidez de la Pitahaya Amarilla en la región Amazonas es superior al 85%, y c) El tiempo de respuesta promedio al utilizar una red neuronal convolucional para determinar la Acidez y Grados Brix de la Pitahaya Amarilla en la región Amazonas es inferior a 30 segundos.

II. MARCO TEÓRICO

Este capítulo describe antecedentes del tema de estudio a nivel nacional e internacional, así como también las bases teóricas que fundamentan la investigación.

De esta manera, **a nivel internacional**, Lozada et al., (2021) llevaron a cabo una investigación aplicada y experimental con el objetivo de analizar el rendimiento de una propuesta de arquitectura basada en Deep Learning y las redes neuronales convolucionales para la detección de la enfermedad del tizón tardío en cultivos de papa. Como base de datos, utilizaron el conjunto PlantVillage, de dónde obtuvieron 1,152 fotografías para realizar el entrenamiento del sistema. Estas imágenes fueron divididas en tres conjuntos, uno para el entrenamiento, con el 85% de las imágenes, otro para la validación, con un 3% de las imágenes y finalmente un conjunto para la puesta a prueba, con el 12% de las imágenes. En total evaluaron 5 arquitecturas diferentes, la propuesta por los autores y una comparación con AlexNet, ZFNet, VGGNet 16 y VGGNet 19. La implementación la realizaron con la biblioteca Tensorflow. Como métricas de evaluación, utilizaron la exactitud, la sensibilidad, la especificidad, la precisión y el puntaje F1, empleando una matriz de confusión para valorar los aciertos verdaderos y falsos (positivos y negativos). Como resultado, la arquitectura propuesta obtuvo excelentes resultados (exactitud 91%, sensibilidad 83%, especificidad 100%, precisión 100%, puntaje F1 90%), solo estando detrás de AlexNet, pero con la ventaja de tener un tamaño mucho menor y la arquitectura profunda. Concluyeron que, por medio de la arquitectura desarrollada, basada en el aprendizaje profundo y redes neuronales, se obtuvo una herramienta valiosa para el diagnóstico o caracterización de cultivos, con resultados precisos, que justifican la creación de aplicaciones finales con interfaces humano máquinas, para ser puestas a prueba en situaciones reales. Así mismo puntualizan que aumentar el conjunto de datos o el número de clases permitiría al modelo obtener una mejor abstracción de las características genéricas (8).

MacEachern et al. (2022) realizaron un estudio planteando como objetivo desarrollar seis modelos de redes neuronales artificiales de aprendizaje profundo para detectar la etapa de madurez en arándanos silvestres. Partiendo de un total de 6,766 imágenes obtenidas en los campos de cosecha, se seleccionaron

aleatoriamente 70% para el entrenamiento y el resto para la comprobación del sistema. Así mismo, las redes fueron entrenadas con 10,000 iteraciones con un factor inicial de 0.0001, comparando las predicciones obtenidas con una matriz de confusión. Entre los resultados, la red YOLOv4 obtuvo la mejor calificación, medida con el indicador F1, en 82%, con un tiempo promedio de respuesta de 7.8 ms y un uso de memoria promedio de 1.63 GB al tratar imágenes de 1280 x 736 píxeles. A pesar de que la tasa de error presentada fue de 24.1%, la data encontrada sobre las labores de campo reales indica que tal desempeño supera los rendimientos obtenidos por medios tradicionales, por lo que se hace factible mejorar las ganancias en los campos de producción (16).

Heredia et al. (2020) realizaron una investigación aplicada, dónde tuvieron como objetivo desarrollar una herramienta para indicar el grado de madurez de las mazorcas de cacao, para lo cual contemplaron el uso de redes neuronales convolucionales. Emplearon un total de 1,992 imágenes para el entrenamiento de la red, en dimensiones de 4033 x 4033 píxeles, 1009 x 1009 píxeles y 505 x 505 píxeles. Cada imagen se clasificó en tres categorías, siendo estas inmaduras, maduras y sobremaduras. De ellas, el 80% fue utilizado en el entrenamiento y 20% en la puesta a prueba. Al respecto, el entrenamiento de la red demoró 68 horas. Paralelamente, el nivel de madurez fue obtenido por medio del juicio de agrónomos expertos en el fruto, quienes clasificaron cada una de las muestras utilizadas, lo cual fue complementado con ensayos de laboratorio dónde se determinó los grados Brix. Sin embargo, en función de los resultados obtenidos, dónde resaltó una precisión de 34.83%, concluyeron que el tamaño de las imágenes utilizadas en la base de datos para el entrenamiento no incide en la eficacia del sistema. Así mismo puntualizan en la importancia de que cada categoría utilizada para el entrenamiento debe tener óptimamente el mismo número de elementos, para así evitar un sobre entrenamiento (17).

Jarrin y Vinueza (2020) tuvieron el objetivo de crear un sistema de pago de frutas empleando el entrenamiento de Redes Neuronales Artificiales Convolucionales. Procedieron con una investigación cuyo diseño fue de carácter experimental, de tipo aplicada, en la cual utilizaron la inteligencia artificial para el procesamiento de imágenes y la identificación de frutas al momento de pago en un supermercado, prescindiendo de la intervención humana, con miras de disminuir

los tiempos de la operación y los errores en la identificación de los productos. En el entrenamiento de la red emplearon las librerías TensorFlow, Keras y OpenCV, además de Transfer Learning, obteniendo una red pre entrenada con más de 1,000 fotografías por cada tipo de fruta de un total de 6 seleccionadas para la investigación, en diferentes ángulos y posiciones. El entrenamiento demoró 106 horas, a partir de lo cual se obtuvo como resultado una precisión en la detección y clasificación de las frutas del 98.9%. Concluyeron que el sistema creado efectivamente facilita el trabajo de los cajeros y el tiempo de compra de los clientes (18).

Moreno (2020), realizó una investigación con el propósito de generar una arquitectura de red basada en redes neuronales convolucionales para la clasificación de Palta Hass según su grado de madurez en ambientes no controlados. Para el entrenamiento de la red empleó 275 fotografías tomadas en campo, así como tres clasificaciones de calidad en función de la presencia y magnitud de defectos en la cáscara de la fruta. Se emplearon 300 ciclos de entrenamiento, con lo cual se obtuvo un 95.73% de precisión. Posteriormente, la prueba del sistema se hizo con 27 imágenes adicionales, lo que arrojó un 80% de sensibilidad, 87.50% de precisión y un indicador F1 de 90.89%. A partir de ello, el autor concluyó que con la aplicación de redes neuronales convolucionales es posible clasificar de forma bastante certera el grado de madurez de la frutal (19).

A nivel nacional, Narciso y Manzano (2021), realizaron una investigación con el objetivo de crear un sistema de visión artificial para ser utilizado en la selección de arándanos, basándose en redes neuronales convolucionales. Se trató de una investigación aplicada de diseño experimental, dónde los autores partieron con la creación de un banco de datos de imágenes para el entrenamiento de la red neuronal en la búsqueda de defectos en la fruta que afectará negativamente el potencial de exportación, en lo cual emplearon el modelo de detección Faster R-CNN con 204 neuronas y 13 capas convoluciones, 13 capas rectificadas y 4 capas de agrupación. Se recurrió a un grupo de expertos para conocer sobre los defectos característicos de la fruta, mientras que los resultados obtenidos fueron analizados empleando el criterio de intersección sobre unión (IoU) para determinar la precisión del sistema. El entrenamiento de la red demandó un periodo aproximado a 5 horas, con 500 imágenes y 41,400 pasos, tras lo cual lograron una tasa de acierto entre

24% a 40.5% para un total de 5 defectos de la fruta considerados, con un tiempo de respuesta de 5 segundos. Concluyeron que tales resultados pueden ser mejorados con la ampliación del banco de imágenes de la base de datos para el entrenamiento (20).

Mosco et al. (2020), realizaron una investigación en la que tuvieron como objetivo lograr la predicción de parámetros de calidad de una muestra de harina de pescado, para lo cual utilizaron una combinación de imágenes hiperespectrales y redes neuronales artificiales. En este sentido, se trató de una investigación aplicada, de tipo experimental, en la que se valieron de una cámara hiperespectrales, la cual hace uso del principio de la espectroscopia del infrarrojo cercano, de manera que pudieron obtener una lectura de los parámetros de la materia prima por métodos no invasivos, lo cual es favorable para no deteriorar la calidad del material, además de ser un método práctico debido a su rapidez. Por su parte, la red neuronal desarrollada con una estructura de red neuronal perceptrón multicapa se empleó para predecir, principalmente, el contenido de humedad en la muestra, lo cual se relaciona con el punto óptimo del proceso de secado de la harina y la calidad del producto final. Como resultado, obtuvieron un sistema que alcanzó una variación promedio del 1.11% entre los resultados reales y los predichos por el sistema, al observar los parámetros de humedad, ceniza y proteína de las muestras. Así mismo, los autores resaltan que el empleo de más capas en la red no es necesariamente favorable, ya que se tiende a producir un sobreajuste (*overfitting*), dónde el modelo se limitaría a aprender los casos particulares del entrenamiento e incapaz de reconocer nuevos valores de entrada. De manera similar, recalcan que una tasa de aprendizaje (*learning rate*) muy bajo haría que la red demore más tiempo en encontrar los valores esperados, mientras que un valor muy alto podría generar que el sistema no converja, aumentando la probabilidad de error (21).

Aguilar y Campoverde (2019) realizaron una investigación en la que abordaron el entrenamiento de una red neuronal convolucional para la clasificación de frutas mediante el procesamiento de imágenes. Fue una investigación aplicada y experimental, dividido en 4 fases principales, la captura de las imágenes para la base de datos, contentiva de 13 categorías de frutas, el procesamiento y etiquetado de las imágenes, el entrenamiento de la red neuronal y la validación del modelo a

partir del porcentaje de eficiencia y el tiempo promedio que demora la respuesta. Para el entrenamiento de la red se empleó el lenguaje Python, la herramienta MobileNetV1 basada en TensorFlow y Matlab. La base de datos estuvo formada por 1,300 imágenes, algunas obtenidas con una cámara digital y otras procedentes de internet. Como resultado, se logró una eficacia del 87%, con tiempos de respuesta de 0.406 segundos. Una prueba adicional reveló que, al reducir las imágenes a 224x224 píxeles, la eficiencia del aprendizaje disminuye. En base a ello concluyeron que el número total de imágenes y la resolución de las mismas son determinantes para la efectividad del aprendizaje (3).

Lazo (2019), llevó a cabo una investigación con el objetivo de utilizar el procesamiento de imágenes por medio de redes neuronales para la detección y clasificación de arándanos en una empresa local. Fue una investigación aplicada de tipo experimental, en las que las imágenes para el entrenamiento fueron obtenidas con el uso de un equipo de espectroscopia infrarrojo cercano (NIR). La base de datos estuvo formada por 1,000 espectros obtenidos de frutos elegidos aleatoriamente en las instalaciones de la empresa. Posteriormente, utilizó la plataforma Keras y TensorFlow. En la etapa de pruebas la aplicación desarrollada se evaluó con una matriz de confusión, donde obtuvo aciertos el 93% para arándanos en buen estado de maduración y sin desperfectos, y el 90% de los casos contrarios, de lo cual se determinó que la aplicación logró una eficacia del 92%. Así mismo, se determinó una exactitud promedio del 97%. Tales resultados permitieron validar la hipótesis de la investigación, dado que, por políticas de la empresa y reglas del mercado, se considera que la cantidad máxima de productos defectuosos por lotes es del 10%, por lo cual, la herramienta desarrollada es capaz de cumplir efectivamente su propósito, reduciendo los costos, los tiempos y optimizando el proceso de selección de la fruta (22).

Cuya y Ramos (2020) llevaron a cabo un estudio aplicado, donde tuvieron el objetivo de evaluar el uso de redes neuronales convolucionales en la elaboración de un sistema de control de calidad para la clasificación de granadilla. Fue una investigación experimental en la que se empleó la metodología de *Design Thinking*. Destacó la aplicación de una encuesta al público consumidor para identificar una serie de factores que inciden sobre la elección de compra de la fruta por parte del consumidor, prevaleciendo el mal aspecto del color, la presencia de golpes y el

tamaño de la fruta, lo que posteriormente sería utilizado para el aprendizaje. La puesta a prueba se llevó a cabo por medio de la clasificación de 435 unidades de fruta, donde la aplicación obtuvo resultados positivos en su evaluación, logrando detectar el 100% de la fruta en buen estado y el 96.79% de la fruta en estado deteriorado, por lo que la efectividad general se calculó en 97.97%. Adicionalmente, los autores concluyeron que el número de filtros aplicados a las imágenes durante el entrenamiento incide en la precisión, donde una menor cantidad de filtros arrojó mejores resultados (23).

Salazar (2020) llevó a cabo un trabajo de investigación con el propósito de diseñar un modelo basado en redes neuronales para la clasificación de Palta Hass. Para el entrenamiento de la red, se tomaron 310 fotografías, a las cuales se les aplicó *Data Augmentation*. Así mismo, el aprendizaje se configuró con un factor de 0.001 y un total de 50 épocas. Cabe destacar que las imágenes para el entrenamiento de la red fueron tomadas en condiciones de luz natural, donde además la resolución de entrada se fijó en 64 x 64 píxeles. Finalmente, se obtuvo una exactitud del 87.5% en la identificación de paltas en buen estado de madurez fisiológica. De esta manera, se comprobó la eficacia del sistema, teniendo la capacidad de disminuir los costos de clasificación de la fruta, dado que actualmente este es un proceso manual que además puede estar sujeto a errores humanos (24).

Respecto a las variables del estudio, se tiene que las **redes neuronales convolucionales** son un subtipo de red neuronal artificial (RNA), sistemas estructurados de cómputo que simulan el funcionamiento natural de las neuronas biológicas, imitando el funcionamiento del propio cerebro humano donde el proceso de aprendizaje o extracción de conocimiento se realiza a partir de un conjunto de datos que, a través del reconocimiento de una serie de patrones, permite identificar tales características en otro conjunto de datos (18). Si bien forman parte de lo denominado como *Machine Learning*, donde se desarrollan modelos de inteligencia artificial que permiten a las computadoras aprender una serie de tareas, las redes neuronales convolucionales (CNN), como método de aprendizaje profundo, han recibido mucha atención durante la última década específicamente en el área dedicada al tratamiento de datos de imágenes (25). Actualmente, este tipo de herramientas son capaces de procesar una gran cantidad de imágenes a color y en gran tamaño en un menor tiempo y con menos recursos de cómputo que sus

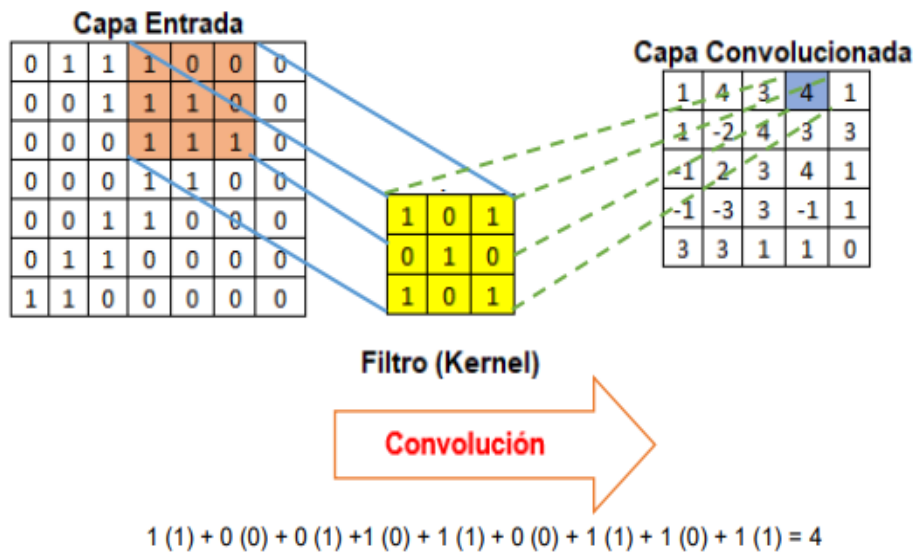
antecedentes (18). Debido a estas características, las redes neuronales convolucionales se han utilizado efectivamente para la organización de galerías de imágenes, el reconocimiento de objetos y personas y particularmente en el reconocimiento facial (22). En cuanto a su funcionamiento, se establecen dos **métodos de aprendizaje** (18):

- Aprendizaje supervisado: en el que se presentan los datos de salida que la red debe ser capaz de generar a partir de señales de entrada. Es decir, produce una función a partir de un conjunto de ejemplos etiquetados para luego hacer una relación con las salidas deseadas.
- Aprendizaje no supervisado: en el que la red neuronal realiza la identificación de las señales de entrada mediante la interconexión entre neuronas. En este sentido, no se produce función alguna, sino que aumenta el número de datos que da origen a la agrupación de información a partir de su similitud (*Clustering*).

En cuanto a su **estructura**, se señalan las siguientes capas (18):

- Capa de entrada: se refiere al conjunto de información que es utilizado para el entrenamiento de la red, que en este caso está conformado por un banco de imágenes. Es una práctica común separar la base de datos en dos grupos, el primero utilizado para el entrenamiento y el segundo para la etapa de validación. Estas imágenes son procesadas como una matriz de datos normalizados.
- Capa de convolución: Es la capa encargada del trabajo de cómputo de alto nivel. En ella, se realizan operaciones matemáticas entre dos funciones cuyo resultado es utilizado para formar una tercera función. De esta manera, el procesamiento de las imágenes emplea valores discretos que ejecutan operaciones sobre una matriz normalizada (*kernel*), cuyo producto escalar entre la matriz y los filtros, generando un mapa de características. Al respecto, es una práctica común utilizar filtros de 2x2 o 3x3. Este proceso se ilustra en la figura 1.

Figura 1. Ejemplo de operación de convolución



Fuente: Tomado de Jarrin y Vinueza, 2020, p. 17.

Terminada la convolución se aplica una función de activación para almacenar únicamente las características deseables de la imagen, como muestra la figura 2.

Figura 2. Aplicación de la función de activación



Fuente: Tomado de Jarrin y Vinueza, 2020, p. 17.

- Capa de agrupamiento: en esta capa se reduce el espacio del tensor de la imagen de entrada, procediendo extraer el máximo valor que representa las características destacadas, como muestra la figura 3.

Figura 3. Ejemplo de capa de agrupamiento



Fuente: Tomado de Jarrin y Vinueza, 2020, p. 18.

- Capa completamente conectada: esta capa se encarga de la conexión entre las neuronas con los nodos de la siguiente capa, realizando una clasificación de las características que fueron extraídas en las convoluciones anteriores.
- Salida: resume las clases para las cuales se llevó a cabo el entrenamiento de la red.

En cuanto a los **métodos de evaluación** del trabajo realizado por las redes neuronales, Lozada et al. (2021) resume los indicadores mostrados en la tabla 1, dónde generalmente la prueba de exactitud es el parámetro más utilizado para la evaluación de este tipo de sistemas (8).

Tabla 1. Métricas de evaluación de resultados

Métrica de evaluación	Fórmula	Descripción
Exactitud	$\frac{TP + TN}{TP + TN + FN + FP}$	Evalúa la cantidad total de instancias que fueron evaluadas e identificadas correctamente
Sensibilidad	$\frac{TP}{TP + FN}$	Evalúa la cantidad de etiquetados correctos, es decir, el número de instancias positivas que fueron clasificadas correctamente
Especificidad	$\frac{TN}{TN + FP}$	Es una evaluación de los valores contrarios a los utilizados en la prueba de sensibilidad, es decir, la cantidad de negativos y falsos negativos identificados
Precisión	$\frac{TP}{TP + FP}$	Evalúa la relación entre los resultados obtenidos y la cantidad de ellos que verdaderamente eran positivos
Puntaje F1	$\frac{2TP}{2TP + FN + FP}$	Evalúa la relación entre la precisión y la sensibilidad.

Nota. TP: totales positivos, TN: totales negativos, FP: falsos positivos, FN: falsos negativos. Tomado de Lozada et al., 2021, p. 4.

Respecto a la **clasificación automatizada de frutas a partir de imágenes**, distintas investigaciones han explorado el uso de redes neuronales, considerando que, el proceso humano es propenso al error, además de invertir una mayor cantidad de tiempo, factores donde la tecnología disponible actualmente puede ser de gran utilidad para el sector agricultor, permitiendo un aumento un aumento de la productividad debido a la eficacia y eficiencia relativa al tiempo y los recursos empleados en el proceso, además de elevar el valor comercial de la producción al disponer de especímenes seleccionados de alta calidad (26).

Por estas razones, el uso de herramientas tecnológicas se ha hecho un nicho en la industria agrícola, impactando positivamente en el crecimiento económico de países como la India o Israel (27).

Entre las técnicas empleadas para la obtención de los parámetros de entrada, se encuentra el uso de la espectroscopia del infrarrojo cercano (NIR), una potente herramienta para la investigación científica que permite la identificación y

el análisis cuantitativo de compuestos químicos, de los cuales se extraen un conjunto de perfiles espectrales que pueden ser analizados individualmente. Este método es especialmente útil al ser de tipo no destructivo, además de no requerir agentes o reactivos químicos para su empleo, por lo que el costo y dificultad de aplicación se ve sustancialmente reducido en comparación a métodos tradicionales como la espectrometría de masa y las técnicas de cromatografía líquida (28, 29). El objetivo principal de su empleo es el establecimiento de modelos usados para predecir o determinar atributos nutritivos o de calidad en muestras, a partir del análisis de los espectros obtenidos (30)

Por su parte, en relación a la **Pitahaya Amarilla (*Hylocereus megalanthus*)**, este es un fruto en forma de baya ovoide, cuya cascara es de color amarillo y está cubierta de espinas. Su interior es de color blanco, presentando una gran cantidad de semillas que son comestibles (figura 4). En cuanto al sabor de la fruta, suele tender hacia lo dulce o agridulce en algunos casos (11). Como alimento, presenta un valor nutricional muy bien valorado, donde destaca la cantidad de vitamina C que contiene por gramo de pulpa (de 4 a 25 mg por cada 100 g de fruta)

Figura 4. Ejemplo de Pitahaya Amarilla



Nota: A: Fruta entera, B: corte transversal. Tomado de Yóplac et al., 2021, p. 12.

Según (11), el **índice de maduración** de la fruta puede ser medido a través de las características fisiológicas que ésta exhibe en su interior o exterior. Entre los primeros, se tienen como indicadores principales la producción de etileno, la firmeza, la cantidad de sólidos solubles, la acidez titulable y el pH, sin embargo,

resalta la necesidad de realizar en todos estos casos ensayos de laboratorio especializados, los cuales usualmente representan un alto costo, además de demorar un tiempo relativamente alto para la obtención de los resultados. En cuanto a las características externas, existen dos indicadores reconocibles, siendo el tamaño nominal de la fruta y su coloración. De estos factores, la figura 5 resume algunos datos sobre la evolución de la fruta a través de su proceso de maduración.

Así mismo, (31) afirman que un indicador confiable de la madurez fisiológica de la fruta es el número de días transcurrido desde la fase de antesis (etapa de floración del árbol), lo cual es evidenciado externamente mediante el color, e internamente en el contenido de sólidos solubles, sin embargo, destacan que el uso de esos valores como referencia puede conducir a inexactitudes si no son calibrados adecuadamente, dado que las condiciones ambientales y edafoclimáticas (zonas geográficas dónde la incidencia de factores como la irradiación solar, la temperatura, la humedad o el tipo de suelo son homogéneos) pueden causar variaciones sutiles en los resultados; particularmente reportan que el término de la maduración después de la antesis ha reportado variaciones entre 25 en México, a 32 días en Brasil.

Figura 5. Etapas de la maduración de la Pitahaya Amarilla



Etapa de crecimiento principal 5: desarrollo reproductivo

510	Aparición de yema reproductiva.
511	Comienzo del oleaje reproductivo.
513	Fin del oleaje reproductivo.
514	Comienzo del alargamiento de yemas.
515	Elongación de las yemas reproductivas.
517	Alargamiento del tubo floral.
519	Fin de la extensión del botón floral, formación de una bola hueca.

Etapa de crecimiento principal 6: floración

610	Flores abiertas.
619	Desvanecimiento de flores y fin de la floración.

Etapa de crecimiento principal 7: desarrollo del fruto

711	Fruto, el ovario comienza a hincharse.
713	30% del tamaño final del fruto.
716	60% del tamaño final del fruto.
719	90% del tamaño final del fruto.

Etapa de crecimiento principal 8: maduración del fruto

811	Comienzo de la maduración de la fruta.
813	Desarrollo avanzado del fruto.
815	Madurez avanzada, madurez fisiológica, fruto para recoger.
817	Fruto maduro.
819	Fruto sobre maduro.

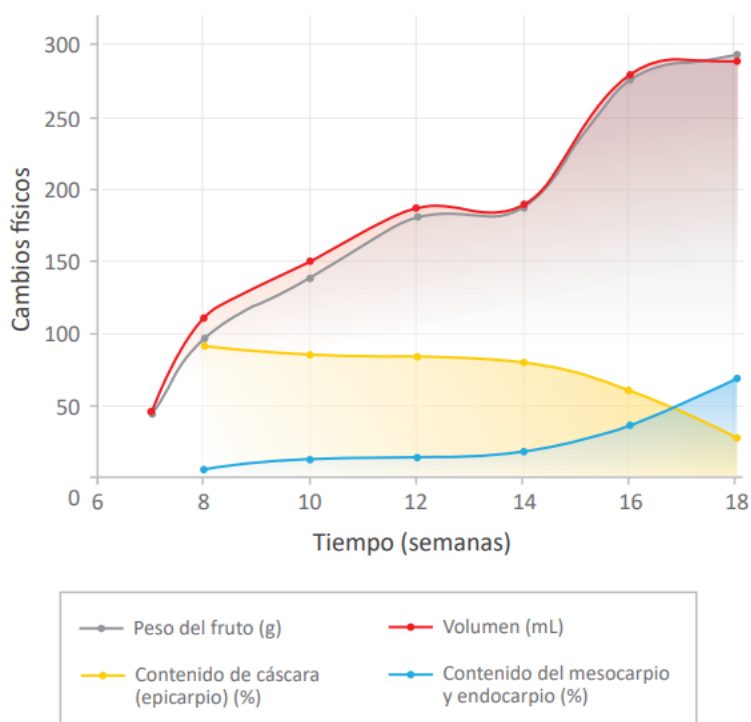
Fuente: Tomado de Yóplac et al., 2021, p. 18.

De lo anterior, se pueden resumir los siguientes indicadores claves:

- El periodo de maduración de la fruta abarca 19 semanas aproximadamente
- El crecimiento de la longitud de la fruta es acelerado las primeras 10 semanas. Para la semana 18 suele alcanzar un valor de 86.8 ± 7.9 mm
- En cuanto al diámetro, puede alcanzar hasta 71.3 ± 5.4 mm en la semana 18
- Puede alcanzar un peso de hasta 293.6 gramos, y un volumen de 289.3 mL.
- El pH suele variar entre 4.8 y 4.6 entre las semanas 8 y 18.
- Los sólidos solubles totales (SST) alcanzan valores de 5.6 a 18.8 grados Bx entre las semanas 12 y 18.

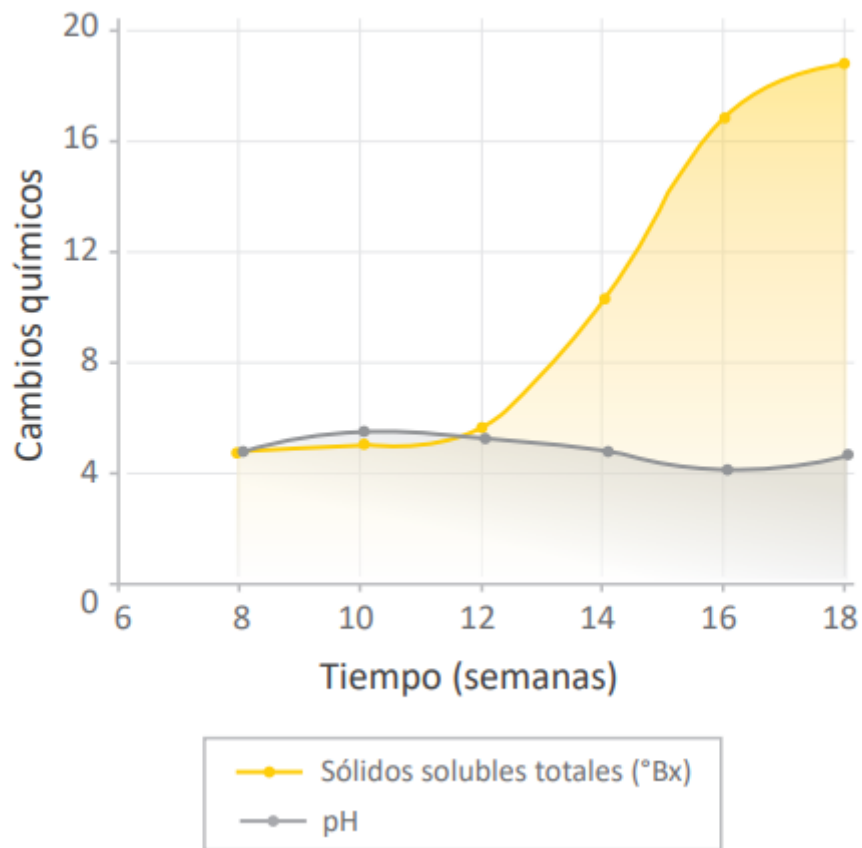
El comportamiento de estos indicadores respecto al tiempo puede visualizarse en las figuras 6 y 7.

Figura 6. Evolución de los cambios físicos en la Pitahaya Amarilla



Fuente: Tomado de Yóplac et al., 2021, p. 25

Figura 7. Evolución de los cambios químicos en la Pitahaya Amarilla



Fuente: Tomado de Yóplac et al., 2021, p. 27.

III. METODOLOGÍA

3.1. Tipo y diseño de investigación

3.1.1. Tipo de investigación

La investigación es de tipo aplicada, la cual se define como aquella que parte de las bases teóricas para ser puesta en práctica en beneficio del hombre, resolviendo problemas que afectan principalmente la vida productiva de la sociedad (32). Bajo este esquema, el propósito de la investigación fue el de obtener un método alternativo para facilitar la determinación de la acidez y los grados Brix de la Pitahaya Amarilla, pudiendo ser utilizado en beneficio de la sociedad de agricultores de la región Amazonas en el proceso de cosecha de la fruta.

Se utilizó un enfoque cuantitativo, dado que los datos a procesar para el logro de los objetivos de la investigación fueron en esencia de naturaleza numérica. En cuanto al nivel, este fue descriptivo, buscando especificar las tendencias características de la población de estudio (33).

3.1.2. Diseño de investigación

El estudio es de tipo experimental, dónde se dio la manipulación de factores, representados en la variable independiente, para observar el efecto producido en la variable dependiente, siendo además de tipo pre experimental, dado que fue un estudio con una sola medición, por lo que el grado de control se considera como mínimo (34). En este sentido, la configuración y el entrenamiento de la red neuronal condicionó la exactitud de los resultados obtenidos en la fase de evaluación. Este diseño obedece al siguiente diagrama:

G x O

Dónde “G” representa los sujetos de análisis y “O” la observación a realizar.

3.2. Variables y operacionalización

Variable independiente: Red neuronal convolucional, Es un modelo informático que simula la etapa de aprendizaje y procesamiento de la información de la misma forma que funciona el sistema nervioso (Capuñay, 2021, p. 50)

Variable dependiente: Determinación de la Acidez y Grados Brix de la Pitahaya Amarilla, Es el proceso de una fruta que ayuda a determinar el índice de madurez que puede ser medido a través de sus características fisicoquímicas que esta exhibe en su interior o exterior, (11)

Dimensiones: Parámetros de la fruta, Exactitud de resultados, tiempo de respuesta promedio.

De lo antes expuesto nuestra matriz o la operacionalización de variables dónde conceptualizamos su definición y los indicadores a implementar se muestra en el anexo 2.

3.3. Población, muestra y muestreo

3.3.1. Población

La población estuvo formada por las imágenes utilizadas para el estudio y las 84 pitahayas amarillas. A su vez, estas fueron seleccionadas según la escala de nivel de madurez, el cual va desde el 0 al 6, utilizando 12 frutos por nivel y procediendo a tomar cierta cantidad de fotografías para cada uno de ellos. En este sentido, la población de estudio fueron 1,620 imágenes y las unidades de pitahaya, en las que se analizaron los grados Brix (sólidos solubles totales - SST), la acidez (AT) y el nivel de pH.

Las frutas fueron obtenidas en el distrito de Churuja, provincia de Bongará, región Amazonas. De cada una de ellas se tomó un grupo de imágenes que se usó para el entrenamiento y validación de la red neuronal. Se utilizaron los siguientes criterios de selección:

- **Criterios de inclusión:** Pitahayas de la variedad *Hylocereus megalanthus*, enteras, de aspecto fresco y consistencia firme.
- **Criterios de exclusión:** Frutas que presenten defectos o imperfecciones por los que serían descartadas para su comercialización.

El tamaño de la población es elegido según los fines de la investigación, dónde, además, el 100% de dichos elementos participaron en el estudio. Por tal

motivo, no se empleó la selección de una porción o muestra representativa de la población ni la aplicación de una técnica de muestreo (34).

3.3.2. Unidad de análisis

Las unidades de análisis fueron las pitahayas amarillas (*Hylocereus megalanthus*).

3.4. Técnicas e instrumentos de recolección de datos

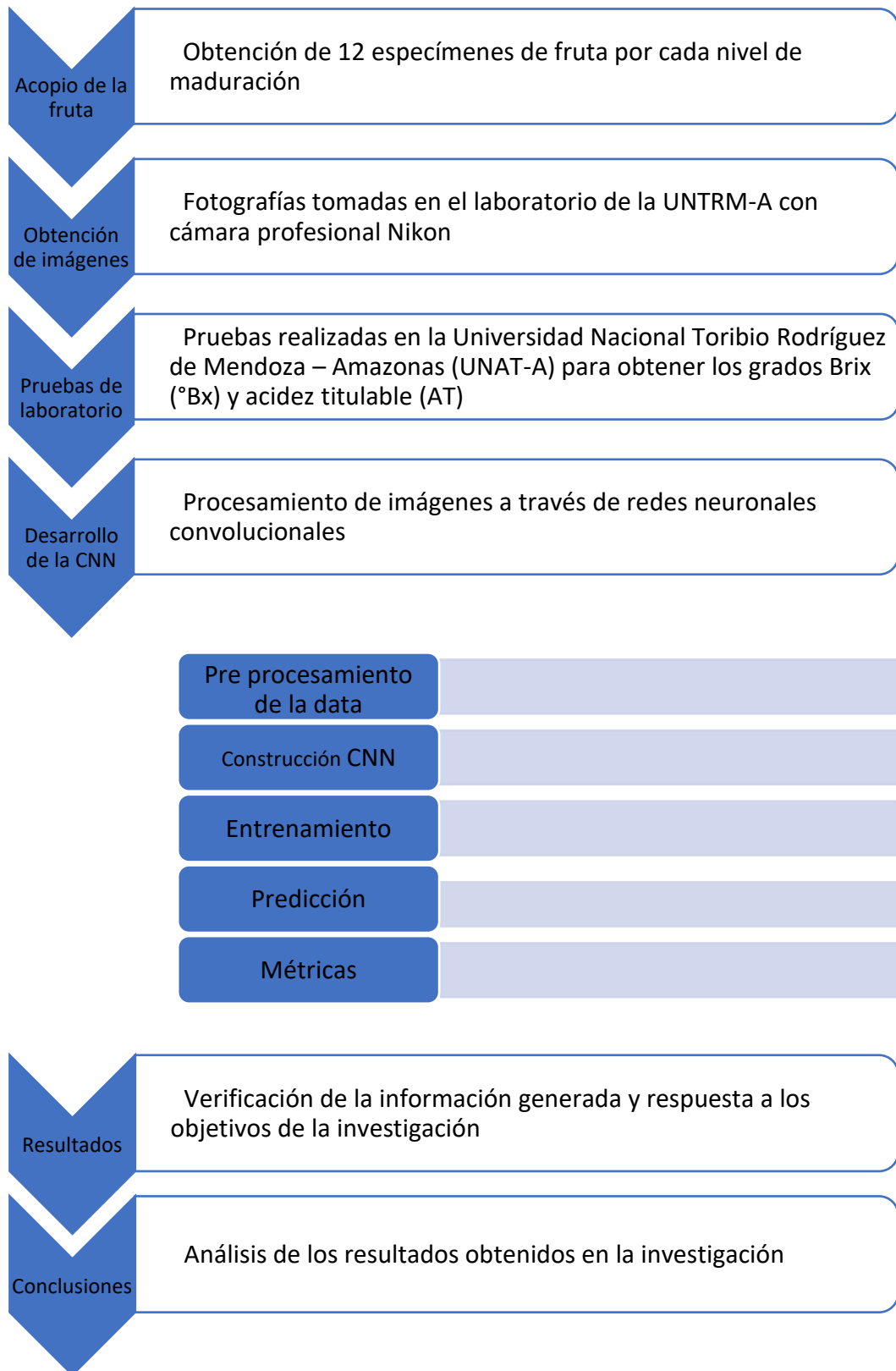
La recolección de datos se realizó aplicando la técnica de la observación directa, la cual se define como el proceso de acopio de información relativa a los sujetos o fenómenos de estudio, dónde el investigador emplea sus sentidos para registrar características, comportamientos, cualidades o propiedades, pudiéndose apoyar en instrumentos de medición (35).

En líneas con lo expresado anteriormente, como instrumentos se utilizaron fichas de registro de observaciones, a fin de poder anotar de manera sistemática y ordenada los datos percibidos del fenómeno o realidad observada. Para la presente investigación se utilizaron dos fichas de observación (anexo 3), en las cuales se registraron, primeramente, las características de las frutas seleccionadas para el estudio, lo que conformó la base de datos a utilizar en el entrenamiento y validación de la red neuronal convolucional. Posteriormente, se empleó la segunda ficha para anotar los resultados obtenidos en la fase de evaluación del sistema, a fin de poder determinar los indicadores relativos a los objetivos propuestos.

3.5. Procedimientos

El desarrollo de una red neuronal convolucional tiene por objetivo identificar y clasificar patrones de algún objeto en estudio, lo cual conlleva una serie de pasos a seguir para permitir y garantizar que los datos con los que la CNN ha sido entrenada brinde y demuestre que sus respuestas son las óptimas. Para esta investigación se detallan los pasos que se realizaron desde el acopio de las pitahayas hasta los resultados obtenidos por la CNN, para ello se muestra en la figura 8 el diagrama del desarrollo.

Figura 8. Diagrama de desarrollo del proyecto



Fuente: Elaboración propia.

3.5.1. Procedimientos en la obtención de imágenes y pruebas de laboratorios

A) Acopio de Pitahaya

Para el desarrollo de este trabajo, en primer lugar, se tuvo que realizar el acopio de la fruta que fue clasificada en función del color que exhibe y el estado de maduración asociado, siguiendo los lineamientos encontrados en la Norma Técnica Colombiana NTC 3554 “Frutas frescas. Pitahaya Amarilla”, dónde se presentan 7 niveles de maduración. De esta manera, se acopió un total de 84 pitahayas amarillas, 12 por cada nivel.

En este proceso se contó con el apoyo del Sr. Santos, productor de Pitahaya Amarilla en el fundo “El Remolino”, perteneciente al mencionado experto, de dónde se recolectaron frutas en los niveles 0, 1 y 2, el día 11 de octubre. Seguidamente, el día 12 de octubre se recolectaron las frutas de los niveles 3, 4, 5, y 6.

B) Obtención de imágenes

Se hicieron dos sesiones de fotos para obtener un total de 1620 fotografías de las frutas en diferentes ángulos de posición. En la primera sesión (11 de octubre de 2022) se fotografiaron las frutas de los niveles 0, 1 y 2, mientras que en la segunda sesión (13 de octubre de 2022) se fotografiaron las frutas de nivel 3, 4, 5 y 6. Se contó con el apoyo del profesor Javier Cabañas López, de la Universidad Nacional Toribio Rodríguez de Mendoza de Amazonas. En el proceso se utilizó una cámara profesional Nikon D5200, a cargo del referido docente. El 80% de las imágenes se usaron para entrenar la red neuronal convolucional y se dejó el 20% para la fase de evaluación de la red.

C) Pruebas de laboratorio

Las pruebas para determinar la acidez titulable (AT), grados Brix (SST) y pH se realizaron en el laboratorio de Bromatología de la UNTRM-A con las 84 frutas de pitahaya que se obtuvieron en el acopio. Se empleó un Refractómetro para determinar los grados Brix (SST), Bureta digital para la acidez titulable y PH Meter para el nivel de pH de cada fruta recolectada (anexo 3), empleando el método de la *Association of Official Analytical Chemists* (AOAC) de 1998 apartado 142.15 según (37). En esta fase se contó con la dirección del docente PhD. Ives Julián Yóplac Tafur docente de la UNTRM-A, con el Ingeniero Julio Santos Peláez, investigador

del Instituto Nacional de Innovación Agraria – Amazonas (INIA), y practicantes de la carrera de agronegocios (figuras 9 y 10).

Figura 9. Clasificación de pitahaya en niveles de maduración 0 al 6



Fuente: Elaboración propia.

Figura 10. Clasificación de pitahaya en niveles de maduración 0 al 6



Fuente: Elaboración propia.

D) Evaluación de los grados Brix

Se procedió a la obtención de jugo de las frutas, picando las pitahayas por la mitad, sacando y exprimiendo la pulpa sobre un vaso de vidrio hasta obtener por lo menos 5 mililitros de jugo (figura 11). Para la evaluación se utilizó un refractómetro de marca HANNA modelo HI96801.

Figura 11. Procedimiento para evaluación de Grados Brix



Fuente: Elaboración propia.

E) Evaluación del pH

Siguiendo la toma de muestras con el procedimiento anterior, para determinar el nivel de pH se utilizó el Ph Meter de marca HANNA modelo Checker,

que consistió en introducir la punta del instrumento dentro del vaso con jugo de 10 ml. hasta obtener la medida de pH, como muestra la figura 12.

Figura 12. Evaluación del pH



Fuente: Elaboración propia.

F) Evaluación de la Acidez Titulable (AT)

Para este procedimiento se utilizó Bureta Digital de marca Titrette clase A, 10 ml de jugo de pitahaya y 150 ml de agua destilada. Se juntaron los líquidos en un vaso de vidrio y se agitó suavemente la mezcla, procediendo a agregar hidróxido de sodio (NaOH) hasta obtener un pH de 8.20 a 8.30 y el resultado obtenido en la bureta digital es el que se registra en la ficha, como muestra la figura 13. Posterior a ello, para determinar la acidez titulable se aplicó la siguiente expresión, según se sugiere en (38):

$$\% \text{ acidez} = B * N * K * 100 / W$$

En Dónde:

B = NaOH consumido en la titulación (mL)

N = Normalidad de NaOH (0,1N)

K = Constante de acidez en la fruta (ácido cítrico) = 0.0642

W = Peso o volumen de la muestra (mL)

Figura 13. *Evaluación de la acidez titulable*



Fuente: Elaboración propia.

También se realizó la tabla de color de acuerdo a la norma técnica de Colombia NTC 3554 en sus diferentes niveles de maduración desde el 0 al 6, como muestra la figura 14.

Figura 14. *Escala de color elaborada según maduración de la fruta*



Fuente: Elaboración propia.

G) Desarrollo de la Red Neuronal Convolutiva

- **Preprocesamiento de la data:** en este paso se realizó la lectura y las importaciones de las librerías que se requiere para procesar la información, dichas librerías son Pandas, numpy, keras, Matplotlib, tensorflow, ImageDataGenerator ver figura 15 y para el desarrollo se utilizó el lenguaje de programación Python en su versión 3.9 y como plataforma el Jupyter notebook

Figura 15. Importación de librerías y preprocesamiento de la data

```

Convolutional Neural Network

Importación de librerías

In [1]: #!pip install tensorflow

In [1]: import tensorflow as tf
        from tensorflow.keras.preprocessing.image import ImageDataGenerator

In [2]: tf.__version__
Out[2]: '2.8.0'

1.- Pre-procesamiento de la data

In [3]: %pylab inline
        import matplotlib.pyplot as plt
        import matplotlib.image as mpimg

        %pylab is deprecated, use %matplotlib inline and import the required libraries.
        Populating the interactive namespace from numpy and matplotlib

In [4]: img = mpimg.imread('dataset/training_set/dogs/dog.343.jpg')
        imgplot = plt.imshow(img)
        plt.show()
    
```

Fuente: Elaboración propia.

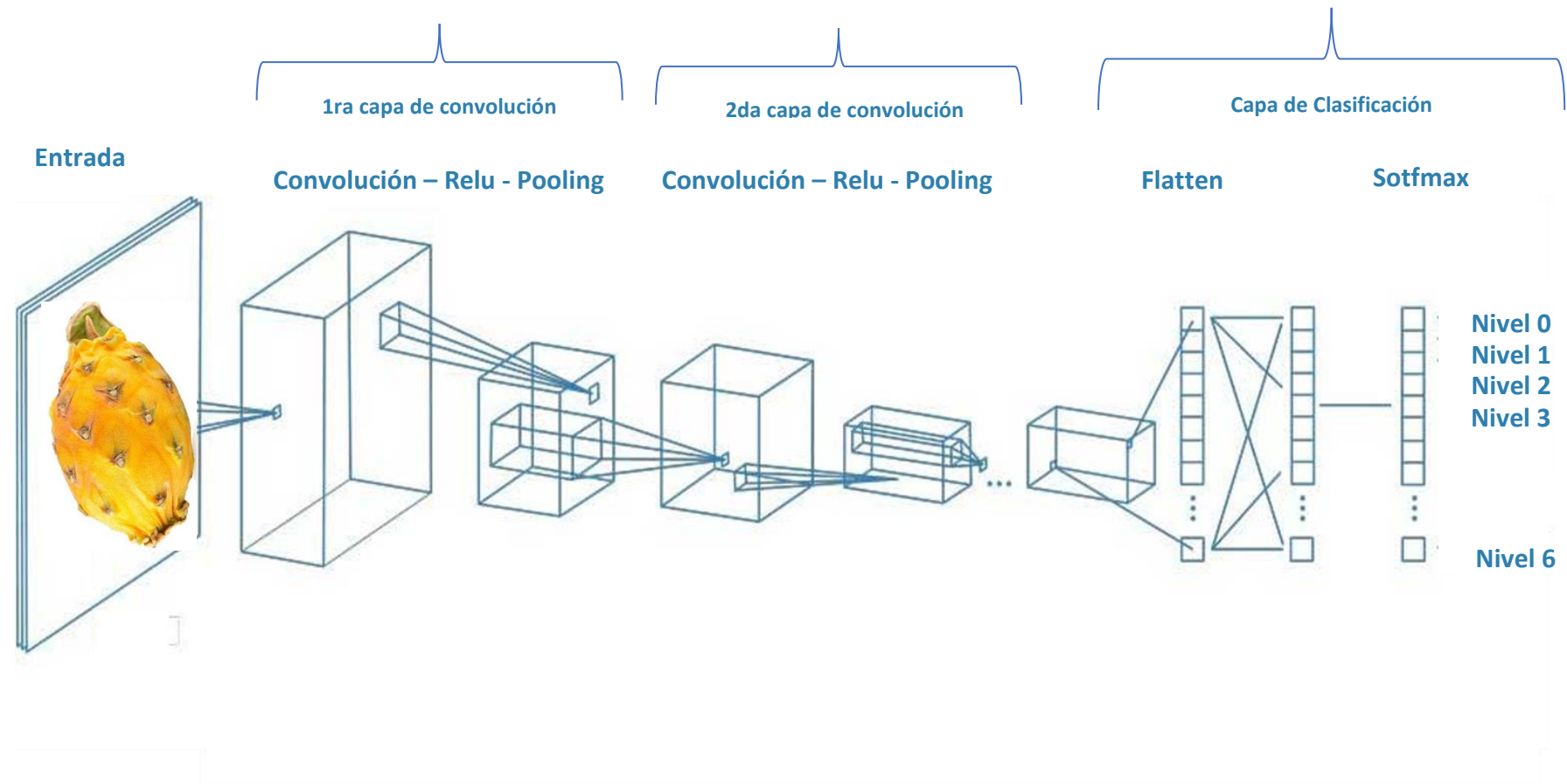
- **Preprocesamiento del set de entrenamiento y del testeo:** después de haber importado las librerías correspondientes, se realizaron las lecturas de imágenes y lo que se debe realizar es normalizar o acotar todas las imágenes a un solo tamaño 160px x 160px y poder generalizar algunos parámetros como Target, color RGB, batch size y se define el modo de las clases, al tener 7 categorías o clases a predecir utilizamos el atributo “categorical” tanto para el set de entrenamiento y para la validación teniendo 1296 imágenes (80% de la población) para entrenar y 324 (20% de la población) imágenes para la validación.

Tabla: Niveles y número de imágenes de pitahaya

Niveles de Pitahaya	Imágenes para Entrenamiento	Imágenes para evaluaciones
Nivel 0	144	36
Nivel 1	192	48
Nivel 2	192	48
Nivel 3	192	48
Nivel 4	192	48
Nivel 5	192	48
Nivel 6	192	48

- **Construcción de la Red Neuronal Convolutiva:** en este procedimiento creamos nuestro modelo de red neuronal el cual estuvo construido por dos capas de convolución, 2 filtros de Max Pooling, 01 capa flaten, 01 capa dropout y 02 capas Densa (activación relu y activación softmax) que indicó que tanta probabilidad tiene cada clase para ellos mostramos la arquitectura desarrollada para nuestro trabajo de investigación

ARQUITECTURA O MODELO DE LA CNN



Fuente: <https://es.mathworks.com/discovery/convolutional-neural-network-matlab.html>

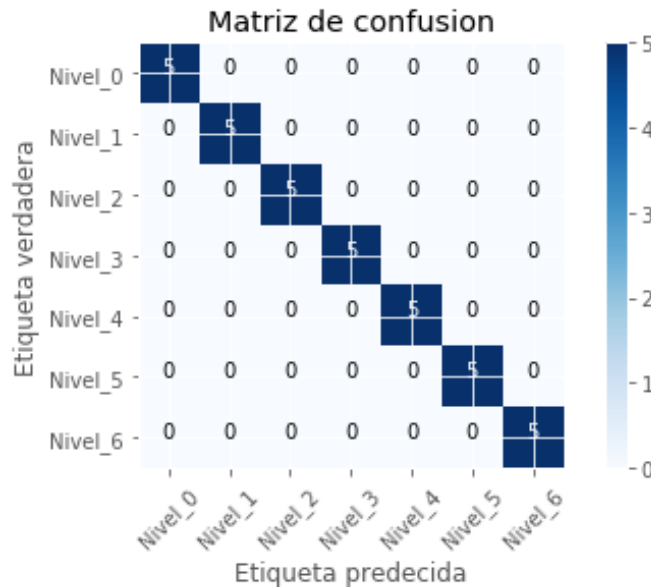
- **Entrenando la Red Neuronal Convolutiva:** para este proceso es necesario llevar y realizar la predicción en base a nuestra experiencia (fallo y error) es la parte más laboriosa realizar y ejecutar nuestro modelo a pruebas con ciertas épocas de entrenamiento para lograr una predicción muy alta, al inicio elaboramos un entrenamiento con 25 épocas teniendo una duración de aproximadamente 23 horas y con unos resultados del 91% para ello reajustamos en el modelo de construcción una capa denominada DropOut que sirve para realizar un salto al modelo cuando entre a una fase de memorización y no de aprendizaje lo que ayudaría a una mejor predicción, generamos la compilación al modelo creado y agregamos los optimizadores, para este modelo usamos Adam, también usamos loss para ver la función de pérdidas para obtener las matrices y accuracy como métricas para visualizar cómo va el entrenamiento en base a los batches realizados para este modelo utilizamos 16 batches que multiplicado a 81 nos da las 1296 imágenes de entrenamiento y finalmente colocamos la cantidad de épocas (16 épocas) a entrenar para tener nuestro resultado final del 0.9961 de exactitud y una pérdida (loss) del 0.0066. es decir, casi 100% de exactitud, ver la figura que se encuentra en el anexo 4, además parte de la buena predicción se debe al balanceo que existe en las imágenes 240 imágenes por cada nivel de color de pitahaya

- **Predicción de la Red Neuronal Convolutiva:** para este proceso se generó una serie de operaciones con Python, cargamos la librería ImageDataGenerator para el procesamiento, redimensionar las imágenes a un tamaño de 160 x 160 para mostrarlas en navegador, bajo su color real RGB, para esta prueba creamos la carpeta de pruebas y colocamos al azar 35 imágenes de las validaciones, también como es parte de nuestros objetivos saber en cuánto tiempo se realiza la predicción, obtuvimos como resultado por las 35 imágenes cargadas un tiempo de 29 segundos tal y como lo muestra la figura que se encuentra en el anexo 4

Finalmente, en la evaluación del sistema, se tomó nota de los aciertos positivos o negativos en la predicción de los grados Brix y acidez titulable, en base a lo cual se determinarán los indicadores relativos a los objetivos de la investigación, conduciendo a la discusión de resultados y la elaboración de las conclusiones del caso.

- **Métricas:** Para ello mostramos la matriz de confusión que después de haber entrenado y haber realizado las predicciones obtuvimos una exactitud del 100%, para ello utilizamos 35 imágenes al azar de todo el dataset y ejecutamos los códigos para obtener los resultados arrojados en las figuras siguientes.

Figura. Matriz de confusión



Fuente: elaboración propia

Figura. Resultados de exactitud

```

7]: print(real_class_indices)
print(predicted_class_indices)

[0 0 0 0 0 1 1 1 1 1 2 2 2 2 2 3 3 3 3 3 4 4 4 4 4 5 5 5 5 5 6 6 6 6 6]
[0 0 0 0 0 1 1 1 1 1 2 2 2 2 2 3 3 3 3 3 4 4 4 4 4 5 5 5 5 5 6 6 6 6 6]

8]: from mlxtend.evaluate import scoring

res = scoring(y_target=real_class_indices, y_predicted=predicted_class_indices, metric='error')
print('Error:', np.round(res * 100, 2), "%")

res = scoring(y_target=real_class_indices, y_predicted=predicted_class_indices, metric='accuracy')
print("Exactitud:", np.round(res*100, 2), "%")

res = scoring(y_target=real_class_indices, y_predicted=predicted_class_indices, metric='average per-class accuracy')
print ("Promedio de exactitud por clase:", np.round(res*100, 2), "%")

Error: 0.0 %
Exactitud: 100.0 %
Promedio de exactitud por clase: 100.0 %

```

Fuente: elaboración propia

3.5.2. Limitaciones

Durante la investigación surgieron los siguientes contratiempos, los cuales limitaron la realización de los procedimientos según lo planificado en el proyecto:

- **Clima y fertilización:** debido a las constantes lluvias en la región amazónica durante los meses de abril a junio, hizo que las floraciones de pitahaya se retrasasen por lo que no se pudo contar con pitahayas de todos los niveles de color en una sola parcela, esto causa que las frutas tengan diferentes características como tamaño, color, dulzor, acidez, lo que al momento de realizar las evaluaciones pudimos notar algunos resultados que no estaban acorde a los niveles del color de la fruta. Dicha información fue corroborada con el evaluador y especialista en pitahaya el ing. Julio Cesar Santos Peláez, indicando que para tener resultados más certeros se debería esperar 3 semanas para tener fruta en todos los niveles y realizar las pruebas nuevamente desde el nivel 4 al 6. Esto retrasó el desarrollo de la red neuronal convolucional según el cronograma de actividades anteriormente planificado.
- **Obtención de muestras en frutas inmaduras:** no se pudo obtener una muestra de jugo en todas las frutas del nivel de maduración 0, llegando a obtener una muestra en solo 3 de las 12 pitahayas. Por lo tanto, es posible que los resultados de pH y acidez titulable presenten un mayor margen de error para este nivel.

3.6. Métodos de análisis de datos

Los resultados a obtener en la validación del sistema fueron analizados utilizando la estadística descriptiva e inferenciales, valiéndose de técnicas numéricas, distribuciones de frecuencia y análisis de varianza que serán mostradas mediante gráficos y cuadros comparativos. En relación a ello, el procesamiento de dichos datos se realizará empleando los programas Excel e Infostat.

3.7. Aspectos éticos

La presente investigación fue desarrollada siguiendo los principios de originalidad, beneficencia, no maleficencia, y aquellas disposiciones que forman

parte del código de ética del 2017 de la Universidad César Vallejo y en su artículo 6 manifestando el respeto a las diferentes investigaciones realizadas por otros investigadores evitando el plagio.

Los datos mostrados son fieles a los obtenidos en los ensayos realizados. Se utilizaron las técnicas de citación apropiadas para reconocer los derechos de autor de las obras citadas, así como a aquellos que colaboraron en la elaboración del estudio.

Se considerará el reconocimiento de los derechos de autor de las referencias bibliográficas utilizadas para fundamentar el caso de estudio y los procedimientos metodológicos para llevar a cabo la investigación. De manera similar, la investigación se hará manteniendo los estándares de responsabilidad y honestidad en el procesamiento de la información.

En cuanto a la obtención de la fruta para realizar los ensayos, se realizará sin perjuicio del medio ambiente, así como también lo relativo a la disposición final del material luego de la investigación.

Para la evaluación de la fruta de la pitahaya en sus aspectos químicos nos basaremos en la Association of Official Analytical Chemists (AOAC) de 1998, donde todas las evaluaciones químicas contienen una serie de procedimientos realizados por otros investigadores (37)

IV. RESULTADOS

4.1. Resultados Descriptivos de la Investigación

4.1.1. Análisis de la Acidez Titulable, Grados Brix y pH

A continuación, se muestran los resultados obtenidos de las evaluaciones de las 12 pitahayas evaluadas en el laboratorio de la UNTRM-A según la tabla de color de 0 a 6 y norma técnica de Colombia NTC-3554 por cada indicador, en base a los resultados obtenidos de los ensayos de laboratorio (ver anexo 3).

- **Indicador 1: Acidez Titulable (AT)**

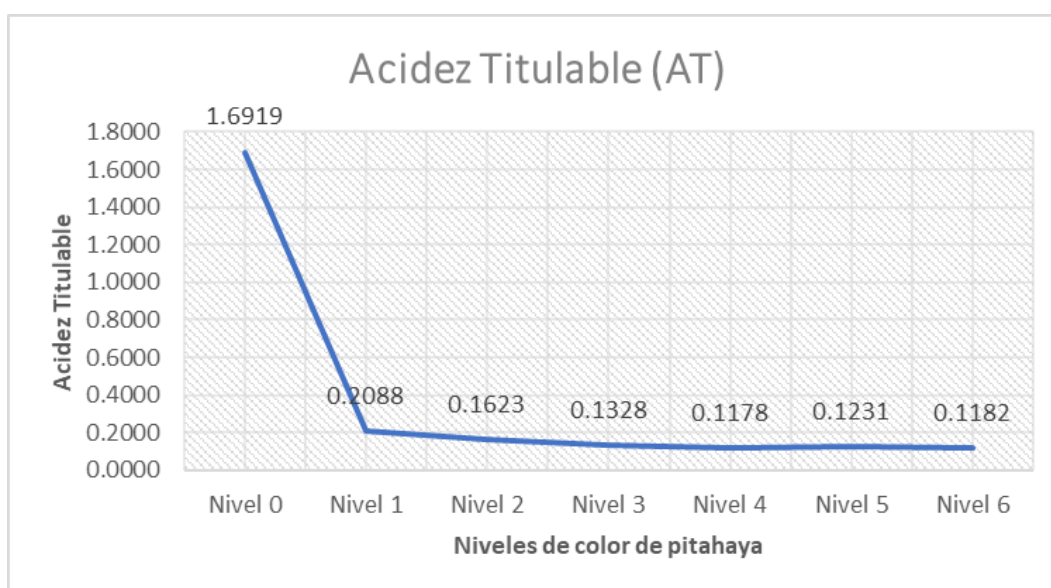
En la tabla 2 y el gráfico de la figura 16 se muestra la frecuencia de las evaluaciones realizadas a 12 pitahayas en los diferentes niveles de color del 0 al 6 para determinar la acidez titulable.

Tabla 2. Tabla de frecuencia Acidez

Nivel	Acidez Titulable (AT) %				
	Valor mínimo	Valor máximo	Promedio	Desviación estándar	Varianza
Nivel 0	1.50	2.01	1.69	0.12	0.0137
Nivel 1	0.19	0.24	0.21	0.02	0.0002
Nivel 2	0.12	0.23	0.16	0.04	0.0016
Nivel 3	0.11	0.22	0.13	0.03	0.0008
Nivel 4	0.09	0.22	0.13	0.03	0.0008
Nivel 5	0.10	0.17	0.12	0.02	0.0004
Nivel 6	0.10	0.15	0.12	0.01	0.0002

Fuente: Elaboración propia.

Figura 16. Evaluación de Acidez Titulable



Fuente: Elaboración propia.

De esta información, se observa que entre el nivel 0 y el nivel 1 ocurre un descenso pronunciado en la acidez de la fruta, mientras que del nivel 1 al nivel 6 el cambio es moderado.

Indicador 2: Grados Brix (°Bx)

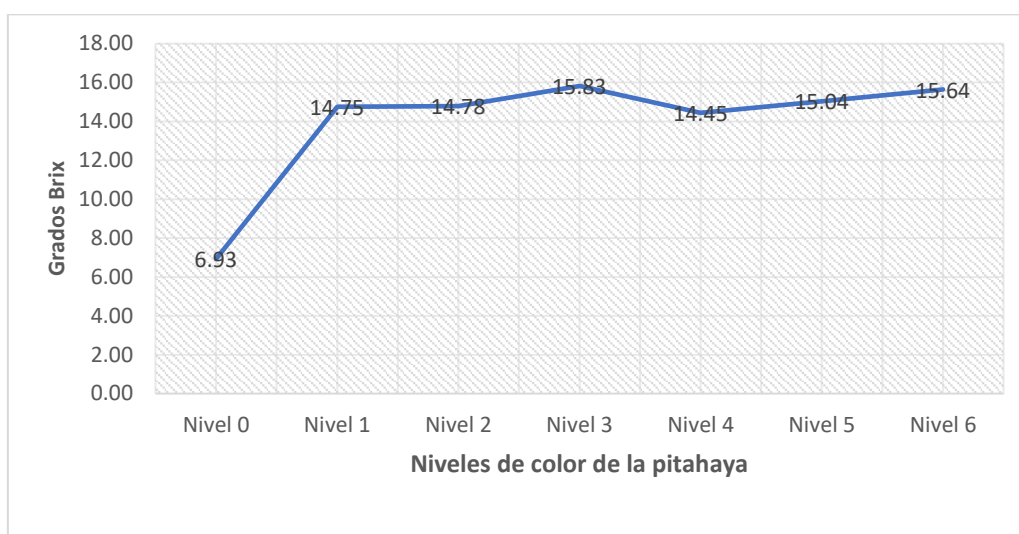
En la tabla 3 y el gráfico de la figura 17 se muestra la frecuencia de las evaluaciones realizadas a las pitahayas en los diferentes niveles de color del 0 al 6 para determinar los grados Brix.

Tabla 3. Tabla de frecuencia Grados Brix

Nivel	Grados Brix				
	Valor mínimo	Valor máximo	Promedio	Desviación estándar	Varianza
Nivel 0	3.60	12.90	6.93	3.14	9.8862
Nivel 1	13.90	15.80	14.75	0.63	0.3955
Nivel 2	13.80	16.40	14.78	0.85	0.7215
Nivel 3	14.20	17.50	15.83	0.99	0.9857
Nivel 4	11.20	17.60	14.45	2.03	4.1009
Nivel 5	11.50	17.00	15.04	1.94	3.7772
Nivel 6	14.80	16.70	15.64	0.58	0.3372

Fuente: Elaboración propia.

Figura 17. Evaluación de grados Brix



Fuente: Elaboración propia.

En lo relativo a este indicador, se observa que entre el nivel 0 y el nivel 1 ocurre un aumento pronunciado de los grados Brix de la fruta, pasando de 6.93 a 14.75 en promedio, lo que representa un incremento de 112%. Llegado a este punto, se observa una tendencia a mantener un valor constante a lo largo de la maduración hasta el nivel 6.

Resultados descriptivos – pH

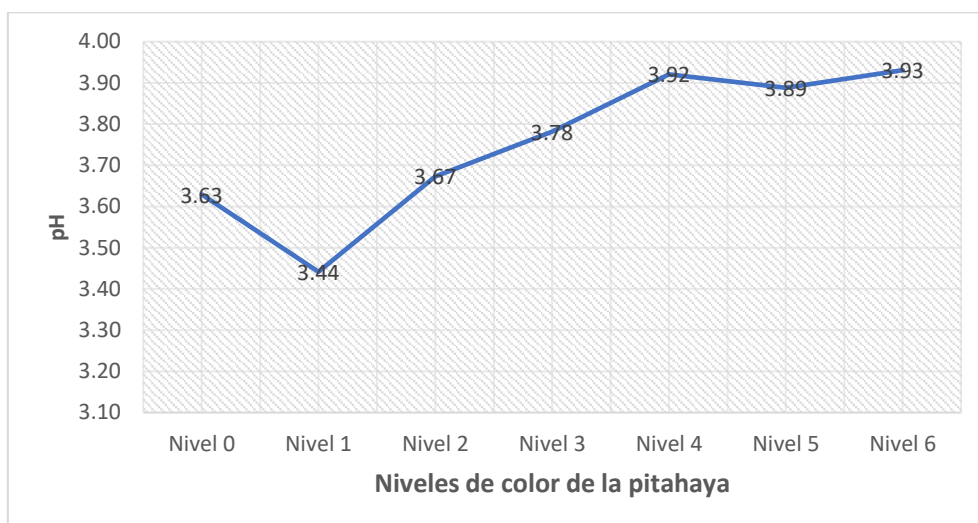
En la siguiente tabla y gráfico mostramos la tabla de frecuencia de las evaluaciones realizadas a 12 pitahayas en los diferentes niveles de color del 0 al 6 para determinar el pH.

Tabla 4. Tabla de frecuencia pH

Nivel	pH				
	Valor mínimo	Valor máximo	Promedio	Desviación estándar	Varianza
Nivel 0	3.35	3.86	3.63	0.11	0.0121
Nivel 1	3.36	3.54	3.44	0.06	0.0032
Nivel 2	3.41	3.93	3.67	0.16	0.0256
Nivel 3	3.52	3.92	3.78	0.12	0.0137
Nivel 4	3.78	4.13	3.92	0.09	0.0086
Nivel 5	3.74	4.01	3.89	0.09	0.0090
Nivel 6	3.75	4.31	3.93	0.15	0.0217

Fuente: Elaboración propia.

Figura 18. Evaluación del pH



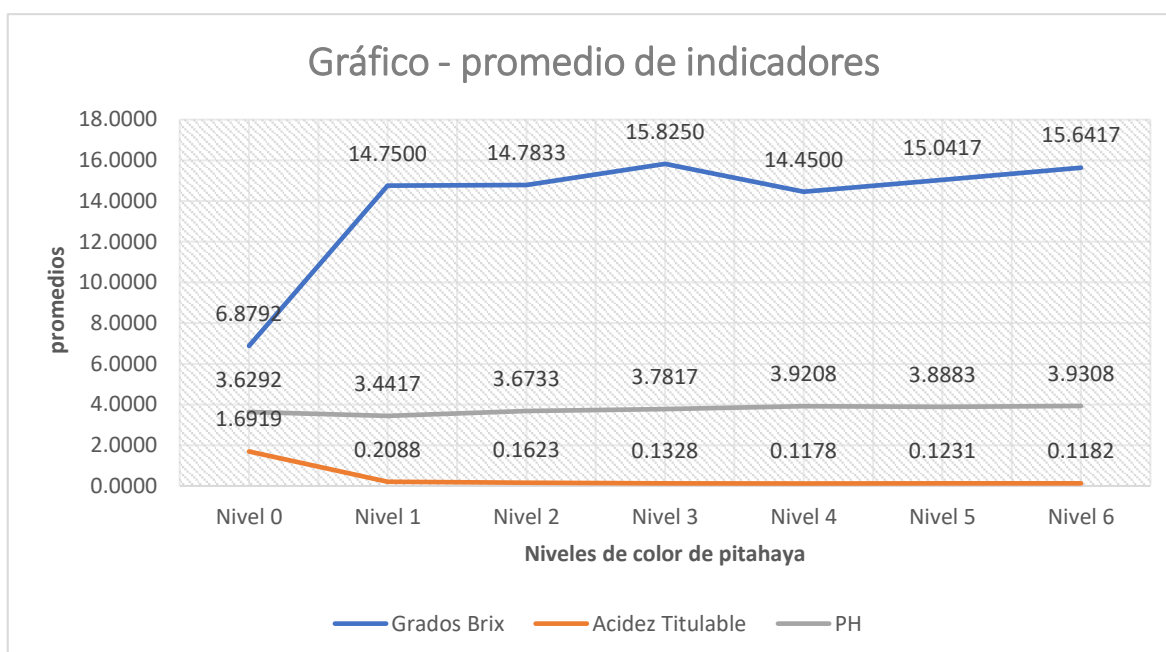
Fuente: Elaboración propia.

En el caso de este indicador, se obtuvo una ligera disminución al pasar del nivel 0 al nivel 1, sin embargo, a partir de ello se observó que el pH de las pitahayas suele incrementarse gradualmente a medida que va madurando.

Resultados descriptivos – Indicadores 1, 2 y 3

En la siguiente figura se muestra un resumen de las evaluaciones realizadas a 12 pitahayas en los diferentes niveles de color del 0 al 6, mostrando los promedios de los indicadores de grados Brix, acidez titulable y pH para determinar el índice de madurez y cómo los indicadores actúan según la pitahaya va cambiando de color desde verde a maduro.

Figura 19. Promedios de Grados Brix (°Bx) – Acidez Titulable (AT) – pH



Fuente: Elaboración propia.

En relación a los resultados mostrados en la figura 19 podemos observar e interpretar lo siguiente:

Con respecto a los grados Brix podemos resolver que conforme avanza los niveles de color de 0 a 6 existe una tendencia a subir, siendo el menor 6.8792 °Bx y el mayor 15.825 ° Bx, es decir cuando la pitahaya tiene un nivel 0 (color verde) su dulzor es bajo, en cambio cuando la fruta tiene un nivel 6 (color amarillo) su dulzor es más agradable o elevado por lo que el color define su estado de dulzor de la pitahaya amarilla

Con respecto a la Acidez Titulable podemos resolver que conforme avanza los niveles de color de 0 a 6 existe una tendencia a bajar, siendo el mayor 1.6919 y el menor 0.1177, es decir cuando la pitahaya se encuentra en un nivel 0 (color verde) su acidez es elevada, en cambio cuando la fruta se encuentra en un nivel 6 (color amarillo) su acidez es baja por lo que el color también define su estado de acidez de la pitahaya amarilla

Con respecto al pH podemos resolver que conforme avanza los niveles de color de 0 a 6 existe una tendencia estable, siendo el mayor 3.9308 y el menor 3.4416, es decir el nivel de PH de la pitahaya amarilla es casi constante en todo el nivel de maduración de la fruta.

Adicional a las evaluaciones realizadas como valor agregado a la investigación decidimos calcular el índice de madurez de pitahaya basándonos en la siguiente fórmula planteada (Velásquez Sofía et. al 2019)

Figura 20. Fórmula para determinar el índice de madurez

DETERMINAR EL ÍNDICE DE MADUREZ

$$I_M = \frac{^{\circ}Brix}{Ac.citrico(Acidez.total)}$$

Contenido de acido cítrico (g) en 1000 ml de jugo

Fuente (37)

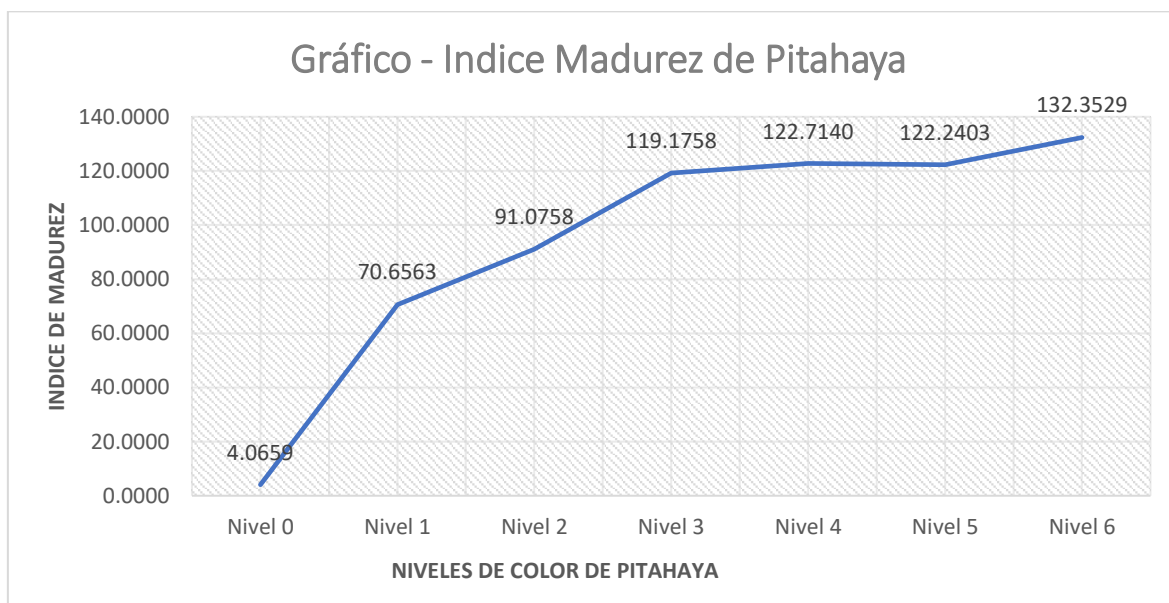
En dónde:

I_M = Índice de Madurez

$^{\circ}Brix$ = Grados Brix

AT = Acidez Titulable

Figura 21. índice de madurez de pitahaya



Fuente: Elaboración propia.

Del gráfico de índice de madurez de pitahaya podemos interpretar que según los niveles de colores van avanzando de 0 a 6 la pitahaya va madurando según su color empezando desde verde hacia amarillo lo que determina que las evaluaciones realizadas son correctas.

4.1.2. Configuración y entrenamiento de la red neuronal

Luego de haber obtenido el set de imágenes que forman la base de datos, se procedió a configurar y entrenar la red neuronal. Los detalles de este proceso se muestran en el anexo 4. En total, se utilizaron 16 épocas de entrenamiento, lo que demoró un total de 3 horas y 28 minutos, con un promedio de 782 segundos por época, tal como muestra la figura 22.

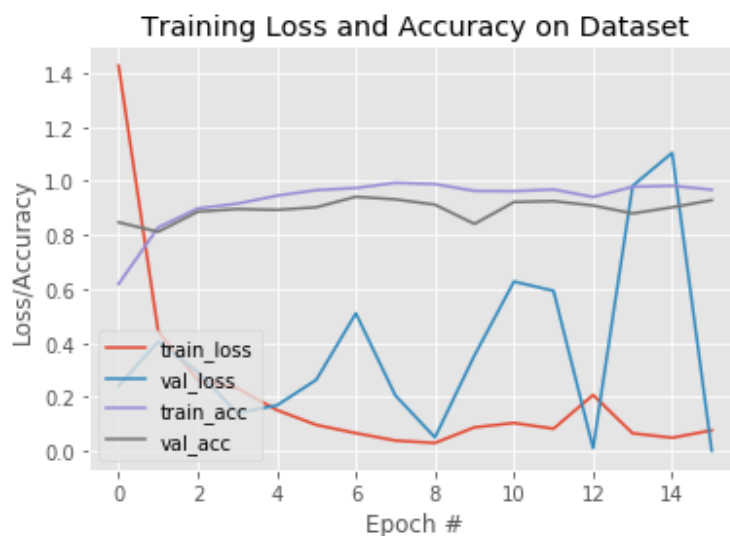
Figura 22. Épocas de entrenamiento de la red

```
Epoch 1/16
81/81 [=====] - 821s 10s/step - loss: 1.4267 - accuracy: 0.6188 - val_loss: 0.2413 - val_accuracy: 0.8469
Epoch 2/16
81/81 [=====] - 819s 10s/step - loss: 0.4439 - accuracy: 0.8279 - val_loss: 0.4052 - val_accuracy: 0.8117
Epoch 3/16
81/81 [=====] - 783s 10s/step - loss: 0.2630 - accuracy: 0.8981 - val_loss: 0.2934 - val_accuracy: 0.8864
Epoch 4/16
81/81 [=====] - 790s 10s/step - loss: 0.2329 - accuracy: 0.9159 - val_loss: 0.1415 - val_accuracy: 0.8961
Epoch 5/16
81/81 [=====] - 792s 10s/step - loss: 0.1520 - accuracy: 0.9452 - val_loss: 0.1688 - val_accuracy: 0.8929
Epoch 6/16
81/81 [=====] - 807s 10s/step - loss: 0.0968 - accuracy: 0.9660 - val_loss: 0.2638 - val_accuracy: 0.9026
Epoch 7/16
81/81 [=====] - 831s 10s/step - loss: 0.0665 - accuracy: 0.9738 - val_loss: 0.5100 - val_accuracy: 0.9416
Epoch 8/16
81/81 [=====] - 791s 10s/step - loss: 0.0388 - accuracy: 0.9931 - val_loss: 0.2060 - val_accuracy: 0.9318
Epoch 9/16
81/81 [=====] - 767s 9s/step - loss: 0.0300 - accuracy: 0.9884 - val_loss: 0.0507 - val_accuracy: 0.9123
Epoch 10/16
81/81 [=====] - 750s 9s/step - loss: 0.0079 - accuracy: 0.9630 - val_loss: 0.3540 - val_accuracy: 0.8409
Epoch 11/16
81/81 [=====] - 768s 9s/step - loss: 0.1040 - accuracy: 0.9622 - val_loss: 0.6276 - val_accuracy: 0.9221
Epoch 12/16
81/81 [=====] - 749s 9s/step - loss: 0.0028 - accuracy: 0.9684 - val_loss: 0.5933 - val_accuracy: 0.9253
Epoch 13/16
81/81 [=====] - 761s 9s/step - loss: 0.2078 - accuracy: 0.9406 - val_loss: 0.0104 - val_accuracy: 0.9091
Epoch 14/16
81/81 [=====] - 768s 9s/step - loss: 0.0659 - accuracy: 0.9784 - val_loss: 0.9818 - val_accuracy: 0.8799
Epoch 15/16
81/81 [=====] - 780s 10s/step - loss: 0.0492 - accuracy: 0.9823 - val_loss: 1.1041 - val_accuracy: 0.9026
Epoch 16/16
81/81 [=====] - 749s 9s/step - loss: 0.0768 - accuracy: 0.9676 - val_loss: 0.0019 - val_accuracy: 0.9286
```

Fuente: elaboración propia.

Por su parte, la figura 23 resume los resultados del entrenamiento, dónde se puede ver que con cada época va aumentando la exactitud de la predicción (*accuracy*).

Figura 23. Resultados del entrenamiento de la red



Fuente: elaboración propia.

Se observa que después de 16 épocas de entrenamiento, la exactitud de la validación (*val_acc*) empieza a mantenerse constante, por lo tanto, se deduce que la cantidad de épocas de entrenamiento fueron suficientes, y no se obtendrían mayores beneficios con épocas adicionales.

4.1.3. Prueba de la red, exactitud y tiempo promedio de respuesta

Para evaluar la exactitud obtenida por la red entrenada, se hizo una prueba con 35 imágenes aleatorias de las que no fueron utilizadas en la fase de entrenamiento, 5 por cada una de los niveles de maduración de la fruta, como muestra la figura 24.

Figura 24. Configuración fase de prueba de la red

```
[ ] test_datagen = ImageDataGenerator(preprocessing_function=preprocess_input)

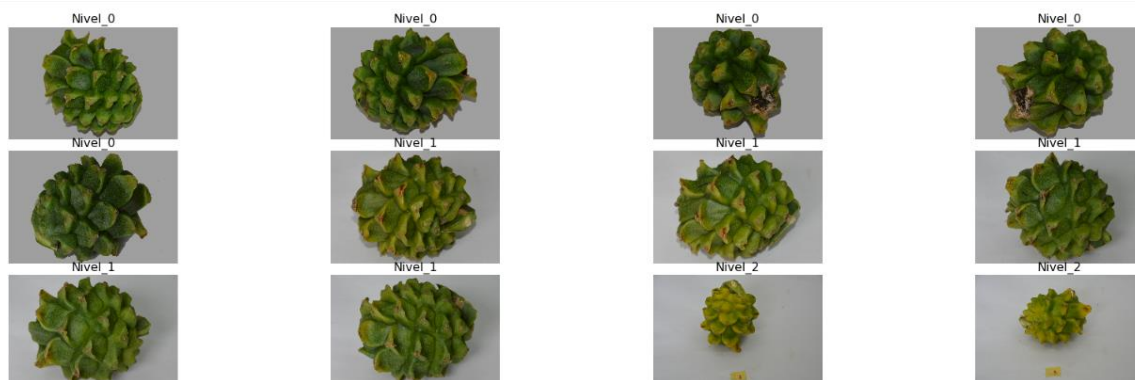
test_generator = test_datagen.flow_from_directory(
    directory="./data/pruebas_expo/",
    target_size=(160, 160),
    color_mode="rgb",
    batch_size=1,
    class_mode=None,
    shuffle=False,
    seed=42
)

Found 35 images belonging to 1 classes.
```

Fuente: elaboración propia.

Como parte de los resultados, se pidió que cada imagen fuera renombrada colocando el nivel de maduración que fue predicho por la red, como muestra la figura 25, dónde se puede observar que en las imágenes utilizadas la fruta estuvo en posiciones y niveles de cercanía aleatorios y los resultados de los indicadores se realizó mediante un archivo xls y csv tal cual lo se muestra en el anexo 4

Figura 24. Impresión de resultados por la red neuronal



Fuente: elaboración propia.

Al respecto, el sistema demoró un total de 29 segundos en hacer la predicción, con un promedio de 828 microsegundos por las 35 fotografías, como muestra la figura 26.

Figura 26. Tiempo empleado en la predicción

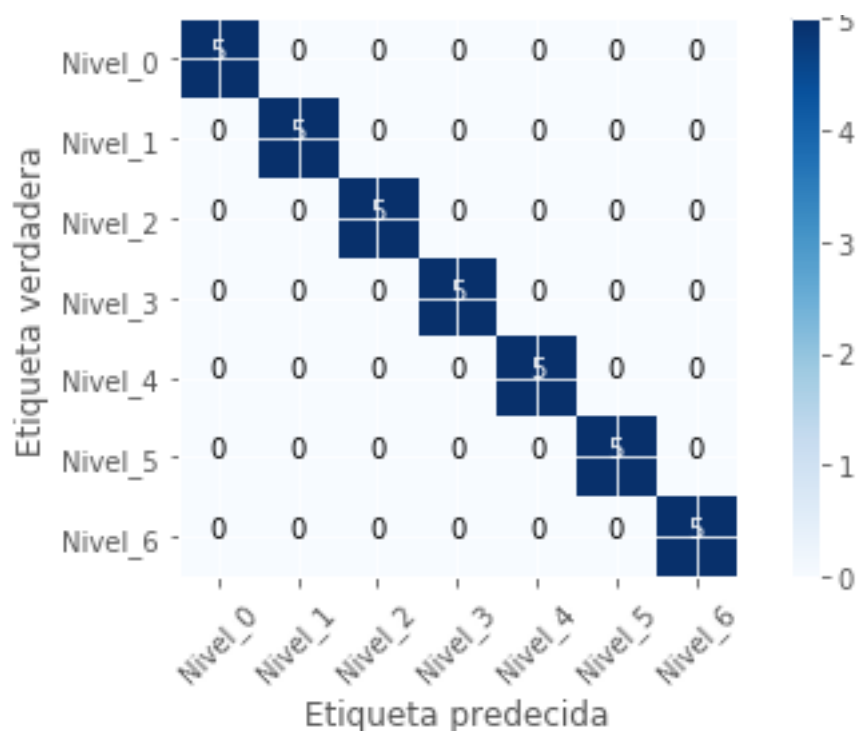
```
STEP_SIZE_TEST=test_generator.n//test_generator.batch_size
test_generator.reset()
pred=model.predict_generator(test_generator, steps=STEP_SIZE_TEST, verbose=1)

35/35 [=====] - 29s 828ms/step
```

Fuente: elaboración propia.

Finalmente, los resultados fueron organizados en una matriz de confusión para poder observar los resultados positivos y negativos de la predicción realizada, la cual se muestra en la figura 27.

Figura 27. Resultados y matriz de confusión obtenida



Fuente: elaboración propia.

En la matriz de confusión, las etiquetas en el eje horizontal corresponden a los valores predichos por la red, y en el eje vertical, los valores verdaderos, determinados a partir de los ensayos de laboratorio realizados anteriormente. Se observa entonces que la red acertó en todos los casos, ubicando cada una de las fotografías en el nivel de madurez de la fruta correspondiente (5 predicciones realizadas que coinciden con los valores reales, en cada nivel de maduración). Al no haber predicciones falsas, se deduce que la exactitud es del 100%, lo que igualmente fue corroborado por el mismo sistema según muestra la figura 28.

Figura 28. Exactitud de las predicciones

```

from mlxtend.evaluate import scoring

res = scoring(y_target=real_class_indices, y_predicted=predicted_class_indices, metric='error')
print('Error:', np.round(res * 100, 2), "%")

res = scoring(y_target=real_class_indices, y_predicted=predicted_class_indices, metric='accuracy')
print("Exactitud:", np.round(res*100, 2), "%")

res = scoring(y_target=real_class_indices, y_predicted=predicted_class_indices, metric='average per-class accuracy')
print ("Promedio de exactitud por clase:", np.round(res*100, 2), "%")

Error: 0.0 %
Exactitud: 100.0 %
Promedio de exactitud por clase: 100.0 %

```

Fuente: elaboración propia.

4.2. Resultados Inferenciales de la Investigación

En relación a la hipótesis planteada en la investigación, según la cual es posible determinar el estado de maduración de la fruta según los parámetros de, grados Brix, Acidez Titulable y pH, se decide utilizar un análisis de comparación de medias ANOVA. Este análisis permite probar hipótesis que están referidas a las medias para determinar si todos los promedios son iguales o diferentes.

4.2.1 Prueba de Bondad:

Se muestra la prueba de normalidad, para determinar la distribución que siguen los datos obtenidos. En ello, se plantean como hipótesis de normalidad:

H_0 = Los datos proceden de una distribución normal

H_a = Los datos no proceden de una distribución normal

Nivel de Significancia:

Confianza: 95%

Significancia (alfa): 5%

Decisión:

- Si p-valor es menor o igual que Alfa, se rechaza la hipótesis nula H_0 y se acepta la hipótesis alterna H_a (los datos no tienen una distribución normal, debiendo emplear pruebas no paramétricas)
- Si p-valor es mayor que Alfa, se acepta la hipótesis nula H_0 y se rechaza la hipótesis alterna H_a (los datos tienen una distribución normal, pudiendo emplear pruebas paramétricas)

Esta revisión se realizó con la prueba estadística Kolmogorov-Smirnov debido a que la población es mayor a 50 elementos (84 pitahayas que fueron evaluadas para determinar los niveles de Acidez titulable, grados Brix y PH). Las figuras 29, 30 y 31 muestran los resultados obtenidos con los datos de la acidez, grados Brix y pH, respectivamente.

Figura 29. Prueba de bondad Kolmogorov – Acidez Titulable

Prueba de bondad de ajuste (Kolmogorov)

Variable	Ajuste	media	varianza	n	Estadístico D	p-valor
AcidezTitulable	Normal(0.2.3)	0.36	0.30	84	0.52	<0.0001

Fuente: elaboración propia. Análisis realizado en el programa InfoStat

Figura 30. Prueba de bondad Kolmogorov – Grados Brix

Prueba de bondad de ajuste (Kolmogorov)

Variable	Ajuste	media	varianza	n	Estadístico D	p-valor
GradosBrix	Normal (0.42.83)	13.91	11.11	84	0.84	<0.0001

Fuente: elaboración propia. Análisis realizado en el programa InfoStat

Figura 31. Prueba de bondad Kolmogorov – pH

Prueba de bondad de ajuste (Kolmogorov)

Variable	Ajuste	media	varianza	n	Estadístico D	p-valor
PH	Normal(0.0.01)	3.75	0.04	84	1.00	<0.0001

Fuente: elaboración propia. Análisis realizado en el programa InfoStat

Ya que en las pruebas se obtuvo que $p < 0.001$, es decir menor al 0.005, se rechaza la hipótesis H_0 y aceptamos la hipótesis alterna H_a , por lo tanto, los datos no proceden de una distribución normal. Se realizó entonces el análisis de varianza (ANOVA).

4.2.2 Análisis de Varianza:

Para esta parte de la investigación todos los datos fueron sometidos a un análisis de varianza (Anova, Anva) el cual permitirá determinar si las medias son iguales o diferentes con una probabilidad de error al 5%.

Análisis de Varianza – Indicador 1: Acidez (AT)

Se plantea la Hipótesis para la acidez titulable:

- H_0 = Todos los niveles de color tienen igual media de acidez titulable
- H_a = Al menos un nivel de Color tiene distinta media de acidez titulable

Figura 32. Análisis de Varianza de Acidez Titulable

Análisis de la varianza

Variable	N	R ²	R ² Aj	CV
Acidez Titulable	84	0.99	0.99	13.58

Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor
Modelo.	24.73	6	4.12	1677.21	<0.0001
NivelColor	24.73	6	4.12	1677.21	<0.0001
Error	0.19	77	2.5E-03		
Total	24.92	83			

Test: Tukey Alfa=0.05 DMS=0.06127

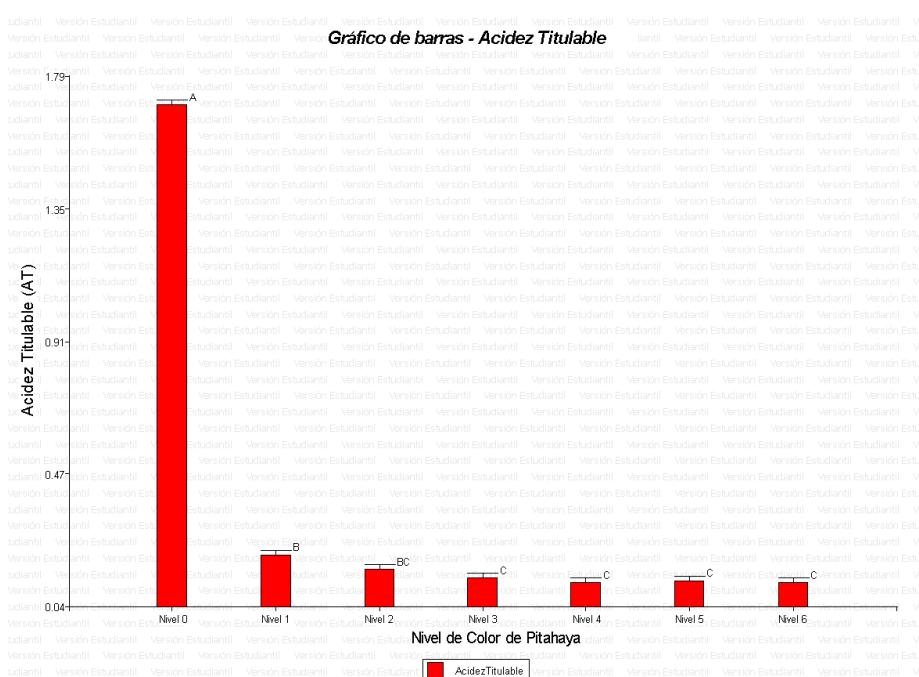
Error: 0.0025 gl: 77

NivelColor	Medias	n	E.E.	
Nivel 0	1.69	12	0.01	A
Nivel 1	0.21	12	0.01	B
Nivel 2	0.16	12	0.01	B C
Nivel 3	0.13	12	0.01	C
Nivel 5	0.12	12	0.01	C
Nivel 6	0.12	12	0.01	C
Nivel 4	0.12	12	0.01	C

Medias con una letra común no son significativamente diferentes (p > 0.05)

Fuente: elaboración propia. Análisis realizado en el programa InfoStat

Figura 33. Gráfico de barras – Acidez Titulable



Fuente: elaboración propia. Análisis realizado en el programa InfoStat

Para esta evaluación se cuenta con un p-valor menor que 0.05, se rechaza la hipótesis nula (Ho) y aceptamos la Hipótesis alterna (Ha): Al menos un nivel de Color tiene distinta media de acidez titulable.

Interpretación:

Finalmente, a través del test de Tukey estadísticamente probamos que las medias de los niveles 4,6,5 y 3 son casi iguales, pero numéricamente son diferentes, el que presenta mayor acidez es el nivel 3.

En cambio, los niveles 0, 1 y 2 sus medias son diferentes al resto de los niveles y se puede determinar que el nivel 0 tiene un valor de 1.69 de acidez y el nivel 6 un valor de 0.12 de acidez, es decir que conforme el nivel de color de la pitahaya sube su acidez tiende a ser más bajo durante su madurez.

Análisis de Varianza – Indicador 2: Grados Brix

Figura 34. Análisis de Varianza de Grados Brix

Análisis de la varianza

Variable	N	R ²	R ² Aj	CV
Grados Brix	84	0.77	0.75	11.95

Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor
Modelo	709.67	6	118.28	42.83	<0.0001
NivelColor	709.67	6	118.28	42.83	<0.0001
Error	212.66	77	2.76		
Total	922.33	83			

Test: Tukey Alfa=0.05 DMS=2.05416

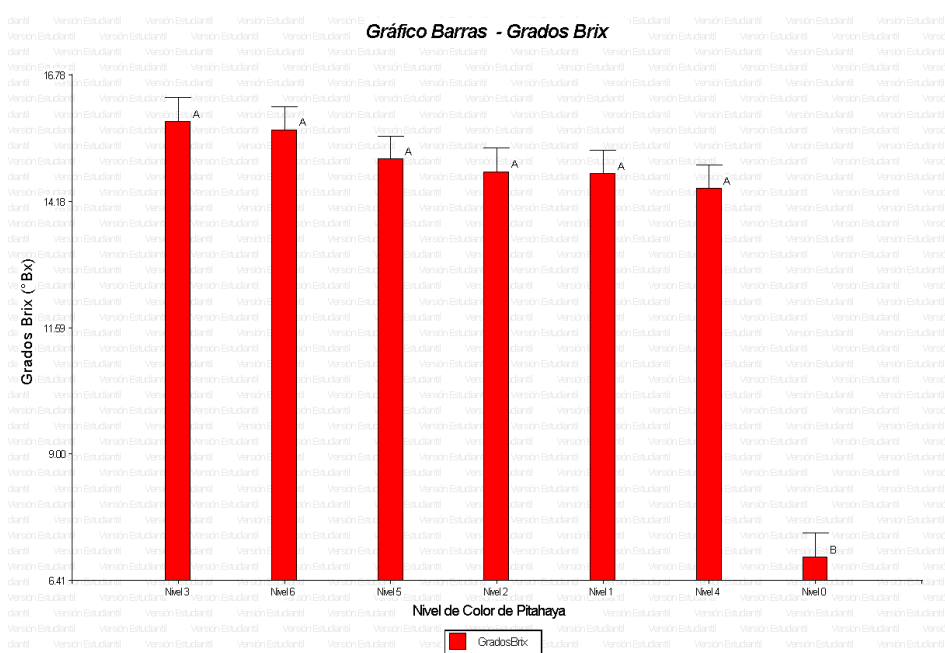
Error: 2.7619 gl: 77

NivelColor	Medias	n	E.E.	
Nivel 3	15.83	12	0.48	A
Nivel 6	15.64	12	0.48	A
Nivel 5	15.04	12	0.48	A
Nivel 2	14.78	12	0.48	A
Nivel 1	14.75	12	0.48	A
Nivel 4	14.45	12	0.48	A
Nivel 0	6.88	12	0.48	B

Medias con una letra común no son significativamente diferentes (p > 0.05)

Fuente: elaboración propia. Análisis realizado en el programa InfoStat

Figura 34. Gráfico de barras – Grados Brix



Fuente: elaboración propia. Análisis realizado en el programa InfoStat

Se plantea la Hipótesis de la investigación para determinar los Grados Brix:

- H_0 = Todos los niveles de color tienen igual media de Grados Brix
- H_a = Al menos un nivel de Color tiene distinta media de Grados Brix

Para esta evaluación contamos con un p-valor de 0.0001 y al ser menor que 0.05, rechazamos la hipótesis nula (H_0) y aceptamos la Hipótesis alterna (H_a): Al menos un nivel de Color tiene distinta media de Grados Brix con respecto a los otros niveles de color.

Interpretación:

Finalmente, a través del test de Tukey estadísticamente probamos que las medias de los niveles 1, 2, 3, 4, 5 y 6 son casi iguales, pero numéricamente son diferentes, el que presenta mayores grados Brix es el nivel 3. En cambio, el nivel 0 su media es diferente al resto de los niveles y se puede determinar que el nivel 0 tiene un valor de 6.88 de grados Brix y el nivel 6 un valor de 15.64 de grados Brix, es decir que conforme el nivel de color de la pitahaya sube su grado Brix tiende a elevarse durante su madurez.

Análisis de Varianza - Indicador 3: pH

Figura 34. Análisis de Varianza de PH

Análisis de la varianza

Variable	N	R ²	R ² Aj	CV
PH	84	0.70	0.67	3.09

Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor
Modelo	2.37	6	0.40	29.46	<0.0001
NivelColor	2.37	6	0.40	29.46	<0.0001
Error	1.03	77	0.01		
Total	3.40	83			

Test: Tukey Alfa=0.05 DMS=0.14314

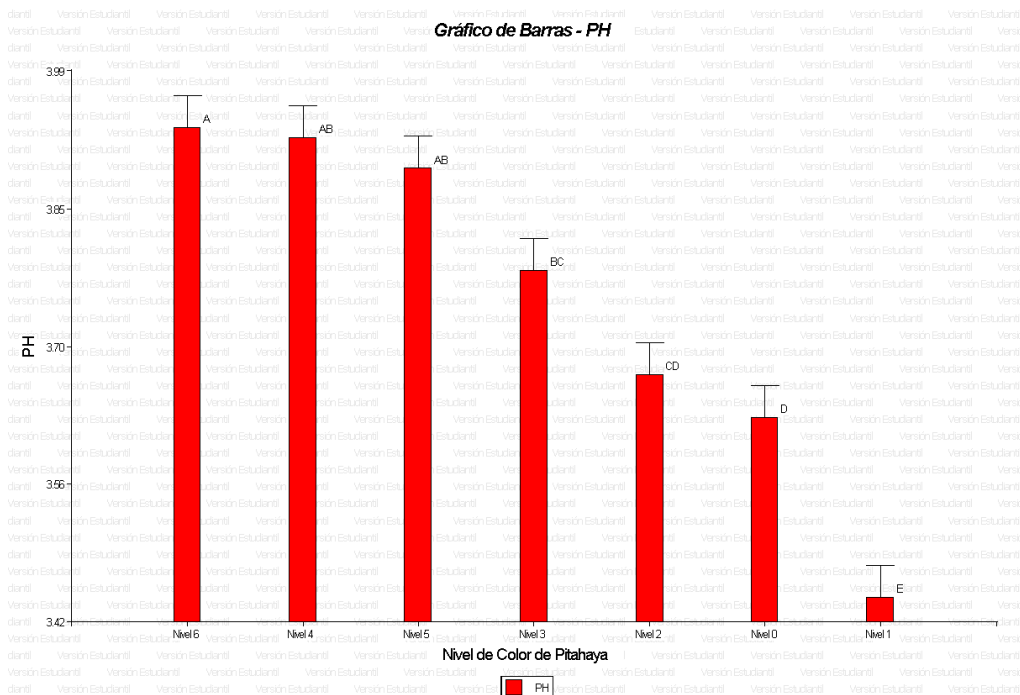
Error: 0.0134 gl: 77

NivelColor	Medias	n	E.E.					
Nivel 6	3.93	12	0.03	A				
Nivel 4	3.92	12	0.03	A	B			
Nivel 5	3.89	12	0.03	A	B			
Nivel 3	3.78	12	0.03		B	C		
Nivel 2	3.67	12	0.03			C	D	
Nivel 0	3.63	12	0.03				D	
Nivel 1	3.44	12	0.03					E

Medias con una letra común no son significativamente diferentes (p > 0.05)

Fuente: elaboración propia. Análisis realizado en el programa InfoStat

Figura 34. Gráfico de barras – PH



Fuente: elaboración propia. Análisis realizado en el programa InfoStat

Se plantea la Hipótesis de la investigación para determinar el PH

Ho = Todos los niveles de color tienen igual media de PH

Ha = Al menos un nivel de Color tiene distinta media de PH

Para esta evaluación contamos con un p-valor de 0.0001 y al ser menor que 0.05, rechazamos la hipótesis nula (Ho) y aceptamos la Hipótesis alterna (Ha), por lo que Al menos un nivel de Color tiene distinta media de pH.

Interpretación:

Finalmente, a través del test de Tukey estadísticamente probamos que las medias de los niveles 4, 5 y 6 son casi iguales al igual que los niveles 3, 4 y 5, pero numéricamente son diferentes, el que presenta mayor PH es el nivel 6.

En cambio, los niveles 0, 1 y 2 sus medias son diferente al resto de los niveles y se puede determinar que el nivel 0 tiene un valor de 3.63 de PH y el nivel 6 un valor de 3.93 de PH, es decir que conforme el nivel de color de la pitahaya sube su nivel de PH es prácticamente constante durante toda su madurez.

Los valores obtenidos en todas las evaluaciones cumplen los requisitos específicos químicos según el Manual de cosecha y postcosecha de pitahaya amarilla (11), tal como se muestra en la figura 34.

Figura 34. Requisitos químicos de la pitahaya amarilla

Características químicas	Color 0 Verde (X ± DE)	Color 1 Pintón (X ± DE)	Color 1 Maduro (X ± DE)	Método de ensayo
pH	4.32 ± 0.12	4.56 ± 0.11	4.73 ± 0.09	Ver 5.2.1.
Sólidos solubles totales (°Bx)	17.92 ± 0.39	18.63 ± 0.58	19.02 ± 0.61	Ver 5.2.2.
Acidez titulable (%)	0.13 ± 0.01	0.12 ± 0.01	0.10 ± 0.01	Ver 5.2.3.
Índice de madurez	140.51 ± 10.80	163.00 ± 16.49	196.78 ± 18.25	Ver 5.2.4.

DE: Desviación estándar.

Fuente: Tomado de Yóplac et al., 2021, p 35.

En el cuadro siguiente podemos encontrar una tabla con los resultados de las hipótesis planteadas en nuestra investigación demostrando que, durante el desarrollo de todo el proyecto, los objetivos e hipótesis fueron aceptados.

Tabla 5: Resultados de Hipótesis

Hipótesis Planteada	Resultados de Hipótesis
H0: Las Redes Neuronales Convolucionales permiten determinar eficientemente la acidez y grados brix de Pitahaya Amarilla en la Región Amazonas	ACEPTADO
H1: Es posible determinar el estado de maduración de la fruta según los parámetros de, grados brix, Acidez y pH	ACEPTADO
H2: El nivel de exactitud al utilizar una red neuronal convolucional para determinar los Grados Brix y Acidez de la Pitahaya Amarilla en la región Amazonas es superior al 85%,	ACEPTADO
H3: El tiempo de respuesta promedio al utilizar una red neuronal convolucional para determinar la acidez y Grados Brix de la Pitahaya Amarilla en la región Amazonas es inferior a 30 segundos.	ACEPTADO

Fuente: Elaboración propia

V. DISCUSIÓN

En la presente investigación se analizaron dos formas de clasificación de las Pitahayas en sus niveles de maduración. La primera, denominado el método tradicional, consistió en la medición de parámetros como los grados Brix, la Acidez Titulable y el pH, mientras que el segundo método se basó en el uso de una red neuronal convolucional y el uso de fotografías, dónde el objetivo fue determinar en qué medida este último resulta eficaz y eficiente, es decir, que proporcionará una respuesta exacta en un bajo periodo de tiempo.

Primeramente, los resultados de laboratorio revelaron que los valores obtenidos de los grados Brix, la Acidez Titulable y el pH no varían de manera significativa entre todos los niveles de maduración de la fruta, siendo muy similares entre algunos niveles sucesivos, lo cual luego fue verificado en el análisis inferencial realizado. En este sentido, se difiere de lo expuesto por Yóplac et al. (2021), pues, así como se ve en la figura 7, los valores de grados Brix deberían variar a medida que la Pitahaya va madurando. El alcance de la investigación no permite establecer mayores conclusiones al respecto, pero resalta lo señalado por el especialista en pitahaya el ing. Julio Cesar Santos Peláez, quien sugiere que, para tener resultados uniformes, las pitahayas deben provenir de una misma parcela, y que sean cosechadas con igualdad de condiciones. En todo caso, se encontró que el índice de madurez, el cual relaciona los grados Brix y la acidez titulable, ofrece mejores resultados, tal como se mostró en la figura 21.

En cuanto al proceso de desarrollo de la red, se hizo notar que una red neuronal convolucional está sujeta a la extracción de ciertas características de un objeto con la finalidad de entender la observación o visualización del ser humano, procediendo para ello en la obtención de una gran cantidad de imágenes para entrenar la red. En el presente trabajo se utilizaron un total de 1620 imágenes, una cantidad similar a la empleada por Lozada et al., (2021), quienes utilizaron 1152 fotografías, aunque estos solo obtuvieron una exactitud en sus predicciones del 91%, mientras que en el presente caso se obtuvo un valor de 100%. Así mismo, se tiene que MacEachern et al. (2022) utilizaron más de 6 mil imágenes, obteniendo también errores en sus predicciones, por lo tanto, se deduce que el número de

imágenes por sí solo no es un factor relevante en la exactitud obtenida por las redes, pues el resultado obtenido demuestra lo contrario.

Por su parte, en relación a los tamaños de imagen utilizados en el entrenamiento, se tiene lo obtenido por Heredia et al. (2020), quienes obtuvieron una precisión de 34.83% empleando imágenes de hasta 4033 x 4033 píxeles, así como también lo hallado por Aguilar y Campoverde (2019), quienes afirmaron que reducir las imágenes disminuye la eficiencia del aprendizaje, no se ha podido llegar a una conclusión definitiva, ya que nuevamente, los resultados obtenidos en el presente investigación contradicen ambos supuestos.

En cuanto al tiempo de respuesta, podemos comparar con la investigación de Narciso y Manzano (2021), quienes tuvieron un tiempo de respuesta en su aplicación de 5 segundos para la selección de arándanos, así como con la de Aguilar y Campoverde (2019), en la que el tiempo fue de 0.406 segundos en la selección de diversas frutas. Habiendo obtenido un tiempo de respuesta de 0.8 segundos, se deduce que el resultado obtenido está dentro de los límites típicamente obtenidos en este tipo de aplicación, y que, además, este tiempo resulta suficiente para trabajar de forma eficiente si se implementa el sistema en una futura aplicación en manos de los productores de pitahaya.

Considerando lo comentado anteriormente, se afirma que el uso de redes neuronales para la selección de frutas y la identificación de características, particularmente el nivel de maduración de las pitahayas, es factible, llegando a obtener resultados satisfactorios en un bajo periodo de tiempo, siendo así más eficiente que la realización de los ensayos de laboratorio tradicionales.

VI. CONCLUSIONES

- Luego de poner a prueba la red neuronal convolucional desarrollada con el objetivo de estimar el nivel de madurez de las pitahayas amarillas, se obtuvieron resultados satisfactorios tanto en la eficacia de la predicción como en la eficiencia de la misma, siendo que para todos los casos de la fase de prueba fueron resueltos con un 100% de exactitud en un bajo periodo de tiempo. Bajo estas consideraciones, este método resulta más efectivo que la realización de ensayos de laboratorios tradicionales, tales que demandan el uso de instrumentos y técnicas especializadas en manos de personal técnico. Siendo que el propósito fundamental de la investigación es la de desarrollar una tecnología que pueda ser puesta en manos de los productores de pitahaya en la región Amazonas, se concluye que la red desarrollada cumple con su propósito, pudiendo ser una herramienta que les permita realizar su trabajo de forma rápida y precisa.
- Al examinar los resultados de los ensayos de laboratorios como se muestran en la figura 19, y luego de aplicar el análisis de varianza, los datos sugieren que no existen diferencias significativas en la acidez, grados Brix y nivel de pH entre las distintas etapas de maduración de la fruta, diferenciándose únicamente en algunas de estas etapas mas no a lo largo de todo el proceso desde el nivel inicial 0 hasta el nivel final 6. En base a ello, se puede concluir que, por sí solos, el contenido de acidez, grados Brix y pH no constituyen un indicador confiable para determinar la madurez de la fruta. Únicamente al calcular el índice de madurez fue posible evidenciar una evolución progresiva del indicador, con una tendencia positiva a medida que la fruta va madurando. Por lo tanto, este último indicador resulta más adecuado para realizar la estimación de la madurez de la pitahaya.
- En cuanto a las métricas de la red neuronal convolucional desarrollada, esta obtuvo una exactitud del 100% en la prueba de validación realizada, como se muestra en la figura 28. Así, el sistema

fue capaz de reconocer en todos los casos el nivel de maduración de la fruta en sus distintas etapas, por lo que se concluye que el sistema es eficaz.

- Al medir el tiempo de respuesta, se obtuvo que las predicciones fueron realizadas en un lapso de 28 segundos para las 35 imágenes utilizadas, es decir, un promedio de 0.8 segundos por elemento. En este sentido, la red es capaz de realizar el trabajo en un periodo de tiempo indiscutiblemente menor que el empleado en la realización de los ensayos de laboratorio tradicionales, lo que se considera como altamente eficiente.
- Finalmente para el entrenamiento de una red neuronal es muy importante contar con una máquina que tenga ciertas características especiales, los cuales ayudarán a que el entrenamiento sea más rápido, de lo dicho anteriormente en esta investigación se usó dos computadoras diferentes, una Core i7 con memoria RAM de 16 Gb y con capacidad de 1 TB de disco duro, teniendo como respuesta al entrenamiento cerca de 4 horas según las figuras del anexo 4, y para el segundo entrenamiento se usó una computadora core i7 con 8 Gb de memoria RAM y 1 TB de disco duro, lo que hizo que el entrenamiento sea más lento teniendo una duración de 26 horas, pero contando con los mismos resultados en las predicciones, lo que tener una máquina de última generación ayudará a disminuir el tiempo de entrenamiento de cualquier red neuronal

VII. RECOMENDACIONES

- Se recomienda proseguir con la creación de una aplicación basada en la red neuronal convolucional aquí desarrollada, para que sea puesta a disposición de los productores de Pitahaya Amarilla en la región de Amazonas.
- A la comunidad académica e investigadores, se recomienda profundizar en el estudio de los indicadores químicos de la Pitahaya Amarilla a lo largo de su maduración, así como indicadores físicos, que puedan contribuir a otras formas de identificación y clasificación de la fruta.
- Se recomienda ampliar el uso de redes neuronales convolucionales como herramientas para la clasificación de frutas u otros productos agrícolas y afines, pudiendo ser de gran ayuda, más eficaces y eficientes que los métodos tradicionales.
- También se puede recomendar para futuras investigaciones solo se evalúen algunos niveles de maduración ya que los resultados arrojan que existe mucha similitud entre los niveles 1-2, los niveles 3-4 y 5-6.
- También recomendamos contar con mayor cantidad de imágenes considerando unas 3000 imágenes para obtener un resultado más certero y que todos los niveles de maduración deben contar con la misma cantidad de imágenes.
- Se recomienda crear aplicaciones móviles o web que sean de acceso al público en general o directa e indirectamente para productores de pitahaya para que les ayude de alguna manera a conocer o determinar los aspectos químicos que tiene esta fruta como la acidez titulable y grados brix.
- Finalmente podemos recomendar el dataset que se requiera utilizar en futuros trabajos de investigación se pueda realizar con imágenes que hayan sido tomados en ambientes no controlables (en campo)

REFERENCIAS

1. *COVID-19 infection prediction from CT scan images of lungs using Iterative Convolution Neural Network Model.* **Madhavi, M y Supraja, D.** 2022, *Advances in Engineering Software*, pág. 103214.
2. *Redes neuronales convolucionales para la clasificación de componentes independientes de rs-fMRI.* **Mera, L y Ochoa, J.** 50, 2021, *TecnoLógicas*, Vol. 24, pág. e1626.
3. **Jarrin, D y Vinueza, D.** *Desarrollo de un sistema prototipo de pago de frutas a través del entrenamiento de redes neuronales artificiales convolucionales.* [Tesis de pregrado, Escuela Politécnica Nacional]. <https://bibdigital.epn.edu.ec/handle/15000/21069> : s.n., 2020.
4. *Convolutional Neural Network applications in additive manufacturing: A review.* **Valizadeh, M y Wolff, S.** 2022, *Advances in Industrial and Manufacturing Engineering*, Vol. 4, pág. 100072.
5. **Lazo, W.** *Espectroscopia con infrarrojo y técnicas de Machine learning y Deep learning para la detección y clasificación de frutas para la agroindustria. Caso: arándanos - Empresa Talsa - 2018.* [Tesis de Maestría, Universidad Privada Antenor Orrego]. <https://repositorio.upao.edu.pe/handle/20.500.12759/5106/> : s.n., 2019.
6. *Classification of Fruits Using Convolutional Neural Network and Transfer Learning Models.* **Pathak, R y Makwana, H.** S3, 2021, *Journal of Management Information and Decision Sciences*, Vol. 24, págs. 1-12.
7. **Hernández, Roberto y Mendoza, Christian.** *Metodología de la Investigación.* Ciudad de Mexico: McGraw-Hill, 2018.
8. **Baena, Guillermina.** *Metodología de la investigación.* Mexico: Grupo Editorial Patria, 2017.
9. **Hernández, Roberto, Fernández, Carlos y Baptista, Pilar.** *Metodología de la investigación.* México: McGraw-Hill, 2014.
10. *Insight: The Key to Faster Progress in Science.* **East, Robert y Ang, Lawrence.** 2021, *Foundations of Science*, Vol. 26, págs. 503-514.

11. **Martínez, R, Palma, A y Velásquez, A.** *Revolución tecnológica e inclusión social.* Santiago : Publicaciones de las Naciones Unidas, 2020.
12. **Yoplac, I, Chávez, R y Santos, J.** *Manual de cosecha y poscosecha de pitahaya amarilla (Hylocereus megalanthus).* Lima : Instituto Nacional de Innovación Agraria.
<https://repositorio.inia.gob.pe/handle/20.500.12955/1420>, 2021.
13. *Pitahaya (Hylocereus spp.): Cultivo, características fisicoquímicas, composición nutricional y compuestos bioactivos.* **Verona, A, Urcia, J y Paucar, L.** 3, 2020, Scientia Agropecuaria, Vol. 11, págs. 439-453.
14. *Sistema de visión artificial basado en redes neuronales convolucionales para la selección de arándanos según estándares de exportación.* **Narciso, W y Manzano, E.** 32, 2021, Campus, Vol. XX VI, págs. 155-166.
15. *Clasificación de frutas basadas en redes neuronales convolucionales.* **Aguilar, J y Campoverde, M.** 1, 2019, Polo del Conocimiento, Vol. 5, págs. 3-22.
16. *Aplicación de redes neuronales convolucionales para la detección del tizón tardío Phytophthora infestans en papa Solanum tuberosum.* **Lozada, W, Suarez, M y Avendaño, E.** 2, 2021, Revista U.D.C.A Actualidad & Divulgación Científica, Vol. 24.
17. *Red neuronal convolucional para la detección de aves exóticas en peligro de extinción.* **Narváez, G.** Universidad de Lima : Fondo Editorial, 2020. Innovando la educación en tecnología. Actas del II Congreso Internacional de Ingeniería de Sistemas. págs. 185-199.
18. *Reconocimiento robusto de lugares mediante redes neuronales convolucionales.* **Lugo, O, Sossa, H y Zamora, E.** 4, 2020, Computación y Sistemas, Vol. 24, págs. 1589-1605.
19. **Cuya, B y Ramos, M.** *Sistema de control de calidad utilizando redes neuronales para la clasificación del estado de la granadilla.* [Tesis de Pregrado, Universidad de Lima].
<https://repositorio.ulima.edu.pe/handle/20.500.12724/13854> : s.n., 2020.
20. **Moreno, J.** *Propuesta de un clasificador del grado de madurez del aguacate hass, mediante una arquitectura de redes neuronales convolucionales.* [Tesis de Maestría, Universidad Tecnológica de Pereira].

<https://repositorio.utp.edu.co/items/8ff5c31e-2142-4cb2-9e92-e44b47bf61d8> : s.n., 2020.

21. **Salazar, J.** *Diseño de un modelo basado en redes neuronales artificiales para la clasificación de palta hass.* [Tesis de Maestría, Pontificia Universidad Católica del Perú]. <https://tesis.pucp.edu.pe/repositorio/handle/20.500.12404/17400> : s.n., 2020.
22. *Determinación de la madurez de mazorcas de Cacao, haciendo uso de redes neuronales convolucionales en un sistema embebido.* **Heredia, J, y otros.** 2, 2020, Revista Colombiana de Computación, Vol. 21, págs. 42-55.
23. *Detection of fruit maturity stage and yield estimation in wild blueberry using deep learning convolutional neural networks.* **MacEachern, C, y otros.** 2022, Smart Agricultural Technology, pág. 100099.
24. **Capuñay, L.** *Modelo basado en redes neuronales para proyectar el rendimiento académico de segundo grado de secundaria en la Institución Educativa N°16093-Jaén.* [Tesis de Maestría, Universidad Cesar Vallejo]. <https://repositorio.ucv.edu.pe/handle/20.500.12692/69442> : s.n., 2021.
25. **Carrasco, S.** *Metodología de la investigación científica.* Lima: Editorial San Marcos, 2017.
26. **Instituto Colombiano de Normas Técnicas y Certificación.** *Norma técnica colombiana NTC 3554 - Frutas Frescas. Pitahaya Amarilla.* 1996.
27. *Determination of the superficial citral content on microparticles: An application of NIR spectroscopy coupled with chemometric tools.* **Yoplac, I, y otros.** 7, 2019, Heliyon, Vol. 5, pág. E02122.
28. *Pitaya fruit quality (Hylocereus undatus [Haworth] Britton & Rose) according to physiological maturity. A review.* **Ortiz, T y Assari, L.** 2020, Revista Colombiana de Ciencias Hortícolas, págs. 63-75.
29. *Detection and Classification of Diseases and Maturity of Dragon Fruits.* **Kulkarni, V, y otros.** 2022, ICT Systems and Sustainability, Vol. 321, págs. 365-374.
30. *Prediction mapping of physicochemical properties in mango by hyperspectral imaging.* **Rungpichayapichet, P, y otros.** 2017, Biosystems Engineering, Vol. 159, págs. 109-120.

31. **Mwiti, D.** Image Classification with Convolutional Neural Networks (CNNs). [En línea] 4 de mayo de 2022. <https://www.kdnuggets.com/2022/05/image-classification-convolutional-neural-networks-cnns.html>.
32. *A Deep Learning-Based Model for Date Fruit Classification.* **Albarrak, K, y otros.** 10, 2022, Sustainability, Vol. 14, págs. 1-16.
33. *Fruit classification quality using convolutional neural network and augmented reality.* **Azani, C, y otros.** 22, 2021, Journal of Theoretical and Applied Information Technology, Vol. 99, págs. 5300-5311.
34. *Dragon fruit: A review of health benefits and nutrients and its sustainable development under climate changes in Vietnam.* **Luu, T, y otros.** 2021, Czech Journal of Food and Science, Vol. 39, págs. 71-94.

ANEXOS

Anexo 1: Matriz de consistencia del proyecto de investigación

Problema	Objetivo	Hipótesis	Variables	Dimensiones	Indicadores	Metodología
<p>General:</p> <p>¿De qué manera se puede se puede utilizar una red neuronal convolucional para determinar la acidez (expresados en ácido cítrico) y Grados Brix (expresado en cantidad °Bx - SST) de la Pitahaya Amarilla en la región Amazonas?</p>	<p>General:</p> <p>Evaluar el uso de Redes Neuronales Convolucionales para determinar la acidez y grados brix de Pitahaya Amarilla en la Región Amazonas</p>	<p>General:</p> <p>Las Redes Neuronales Convolucionales permiten determinar eficientemente la acidez y grados brix de Pitahaya Amarilla en la Región Amazonas</p>	<p>Independiente:</p> <p>Red neuronal convolucional</p> <p>Es un modelo informático que simula la etapa de aprendizaje y procesamiento de la información de la misma forma que funciona el sistema nervioso (Capuñay, 2021, p. 50)</p>	<p>Configuración de la red neuronal</p>		<p>Tipo de investigación:</p> <p>Aplicada - Descriptiva</p> <p>Diseño de la investigación:</p> <p>Experimental</p> <p>Población: 84 Pitahayas Amarillas y 1620 imágenes</p> <p>Muestra: 100% de la población</p> <p>Instrumentos de recolección de datos:</p> <p>Fichas de registro de observación</p> <p>Método de análisis:</p> <p>Estadística descriptiva e inferencial</p>
<p>Específicos:</p> <p>a) ¿Qué parámetros de SST, AT y pH presenta la Pitahaya Amarilla según</p>	<p>Específicos:</p> <p>a) Determinar los parámetros SST, Acidez Titulable y pH de la Pitahaya Amarilla según los</p>	<p>Específicos:</p> <p>a) Es posible determinar el estado de maduración de la fruta según los parámetros de, SST,</p>	<p>Dependiente:</p> <p>Determinación de la Acidez y Grados Brix de la Pitahaya Amarilla</p>	<p>Parámetros de la fruta</p>	<p>Acidez Titulable - AT</p> <p>Grados Brix - SST</p> <p>Ph</p>	

<p>los estados de maduración?</p> <p>b) ¿Qué nivel de exactitud se obtiene al utilizar una red neuronal convolucional para determinar la Acidez y Grados Brix de la Pitahaya Amarilla en la región Amazonas?</p> <p>c) ¿Cuál es el tiempo de respuesta promedio al utilizar una red neuronal convolucional para determinar la Acidez y Grados Brix de la Pitahaya Amarilla en la región Amazonas?</p>	<p>estados de maduración</p> <p>b) Determinar el nivel de exactitud al utilizar una red neuronal convolucional para determinar la Acidez y Grados Brix de la Pitahaya Amarilla en la región Amazonas</p> <p>c) Determinar el tiempo de respuesta promedio al utilizar una red neuronal convolucional para determinar la Acidez y Grados Brix de la Pitahaya Amarilla en la región Amazonas</p>	<p>Acidez Titulable y pH de la Pitahaya Amarilla según los estados de maduración</p> <p>b) El nivel de exactitud al utilizar una red neuronal convolucional para determinar la Acidez y Grados Brix de la Pitahaya Amarilla en la región Amazonas es superior al 85%</p> <p>c) El tiempo de respuesta promedio al utilizar una red neuronal convolucional para determinar la Acidez y Grados Brix de la Pitahaya Amarilla en la región Amazonas es inferior a 30 segundos</p>	<p>Es el proceso de una fruta que ayuda a determinar el índice de madurez que puede ser medido a través de sus características fisicoquímicas que esta exhibe en su interior o exterior, teniendo la necesidad de realizarse en laboratorios especializados (Yóplac, 2021)</p>	<p>Exactitud de los resultados</p> <p>Tiempo de respuesta promedio</p>	<p>Totales positivos</p> <p>Totales negativos</p> <p>Falsos positivos</p> <p>Falsos negativos</p> <p>Tiempo de respuesta</p> <p>Cantidad de evaluaciones</p>	<p>Área de estudio: Región amazonas</p>
---	--	---	--	--	--	--

Anexo 2: Variables y operacionalización

Variables	Definición Conceptual	Definición Operacional	Dimensión	Indicador	Escala de Medición
Independiente: Red neuronal convolucional	Capuñay (2021, p. 50) Alude que las redes neuronales son modelos informáticos que simulan la etapa de aprendizaje y procesamiento de la información de la misma forma que funciona el sistema nervioso de la persona. Se menciona que es el proceso dónde la red neuronal procesa la data suministrada según valores aprendidos en el entrenamiento con la finalidad de entregar información importante de forma automática (Pathak & Makwana, 2021, p. 3)	La red neuronal convolucional que se implementará está basada para determinar las características fisicoquímicas de una pitahaya mediante un modelo de red neuronal que fueron evaluados en un proceso de laboratorio y determinar la exactitud de los resultados arrojados tanto en laboratorio como en el modelo implementado.	Configuración de la red neuronal		De razón
			Entrenamiento de la red		
Dependiente: Determinación la Acidez y Grados Brix de la Pitahaya Amarilla	Es el proceso de una fruta que ayuda a determinar el índice de madurez que puede ser medido a través de sus características fisicoquímicas que esta exhibe en su interior o exterior, (11) teniendo la necesidad de realizarse en laboratorios especializados.	Las evaluaciones o procedimientos de características químicas de frutas lo establece la AOAC 1998 (Association of Official Analytical Chemists) en su apartado 942.15 establece el procedimiento para el método de obtención de titulación es decir para determinar la	Parámetros de la fruta	Acidez Titulable, Grados Brix, Ph	Nominal
			Exactitud de los resultados	Totales positivos Totales negativos Falsos positivos Falsos negativos	De razón

	<p>(31) afirma que un indicador confiable de madurez fisiológica es el color e internamente o químicamente son los SST o grados brix, acidez titulable, Ph entre otros.</p>	<p>acidez de la fruta apoyado por una bureta digital, así mismo mediante un refractómetro se pretende determinar los valores de grados brix y por medio de un Ph Meter el nivel de Ph de la pitahaya, la UNTRM-A tiene los instrumentos necesarios para obtener las características requeridas los cuales serán usados para determinar los valores mencionados</p>	<p>Tiempo de respuesta promedio</p>	<p>Tiempo de respuesta Cantidad de evaluaciones</p>	
--	---	--	-------------------------------------	---	--

Anexo 3. Resultados de los ensayos de laboratorio para determinar los grados Brix, la acidez titulable y el pH



UNIVERSIDAD CÉSAR VALLEJO

Nombre de la investigación:

Redes Neuronales Convolucionales para determinar la acidez y grados brix de Pitahaya Amarilla (*Hylocereus megalanthus*) en la Región Amazonas

Autor:

Cueva Caro, Jeison Enrique

**Ficha de registro
Características de la fruta**

Estado de madurez de la Pitahaya Amarilla según tabla de color

Color O	X	Color 1	Color 2	Color 3	Color 4	Color 5	Color 6
N° Espécimen	Grados Brix (°Bx)	Acidez Total (AT)	pH	N° Imágenes obtenidas			
1	11.4	1.69	3.63	15			
2	4.2	1.58	3.63	15			
3	6.3	1.69	3.63	15			
4	9	2.01	3.86	15			
5	6.3	1.69	3.67	15			
6	12.9	1.5	3.35	15			
7	0	1.69	3.63	15			
8	3.6	1.69	3.63	15			
9	5	1.69	3.63	15			
10	5.4	1.69	3.63	15			
11	8.5	1.69	3.63	15			
12	3.6	1.69	3.63	15			
Total, de imágenes del estado de madurez				180			



Nombre de la investigación:

Redes Neuronales Convolucionales para determinar la acidez y grados brix de Pitahaya Amarilla (*Hylocereus megalanthus*) en la Región Amazonas

Autor:

Cueva Caro, Jeison Enrique

**Ficha de registro
Características de la fruta**

Estado de madurez de la Pitahaya Amarilla según tabla de color

Color 0	Color 1	Color 2	Color 3	X	Color 4	Color 5	Color 6
Nº Especimen	Grados Brix (°Bx)	Acidez Total (AT)	pH	Nº Imágenes obtenidas			
1	16.5	0.13	3.73	20			
2	14.5	0.13	3.72	20			
3	16.1	0.12	3.78	20			
4	17.5	0.22	3.74	20			
5	17	0.13	3.91	20			
6	15.9	0.11	3.92	20			
7	15.2	0.15	3.67	20			
8	14.2	0.11	3.79	20			
9	15.3	0.12	3.92	20			
10	16.5	0.13	3.83	20			
11	16.1	0.13	3.85	20			
12	15.1	0.12	3.52	20			
Total de imágenes del estado de madurez				240			

Anexo 4. Configuración, entrenamiento y prueba de la red neuronal

```
import pandas as pd
import numpy as np
import os
import keras
import matplotlib.pyplot as plt
import tensorflow as tf
from keras import applications
from keras.utils import to_categorical
from keras.layers import Dense, GlobalAveragePooling2D
from keras.layers.convolutional import Conv2D
from keras.layers import Dense, Dropout, Convolution2D, MaxPooling2D
from keras.layers.convolutional import MaxPooling2D
from keras.layers import Flatten
from keras.applications import MobileNet
from keras.preprocessing import image
from keras.applications.mobilenet import preprocess_input
from keras.preprocessing.image import ImageDataGenerator
from keras.models import Model, Sequential
from keras.optimizers import Adam
from mlxtend.evaluate import confusion_matrix
from keras import backend as K
import itertools
```

Using TensorFlow backend.

```
K.clear_session()
```

```
filtrosConv1 = 32 # numero de filtros que aplicaremos tras la primera convolucion
filtrosConv2 = 64 # numero de filtros que aplicaremos tras la segunda convolucion
tamano_filtro1 = (3, 3) #para primera convolucion
tamano_filtro2 = (2, 2) #para segunda convolucion
tamano_pool = (2, 2) # para mejorar el avance de la convolucion
```

```
model = Sequential() #varias capas apiladas entre ellas
model.add(Convolution2D(filtrosConv1, tamano_filtro1, padding='same', input_shape=(160, 160, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=tamano_pool))
```

```
model.add(Convolution2D(filtrosConv2, tamano_filtro2, padding='same'))
model.add(MaxPooling2D(pool_size=tamano_pool))
```

```
model.add(Flatten()) #imagen profunda la vamos a hacer plana, es decir solo una dimension va tener toda nuestra informacion de la cnn
model.add(Dense(256, activation='relu')) #añadimos una capa que va conectar la anterior y la sgte con 256 conexiones
model.add(Dropout(0.5)) #apagamos 50% de las neuronas cada paso, así evitamos sobre ajustar (evitar un solo camino de entrenamiento)
```

```
model.add(Dense(7, activation='softmax')) # ultima capa softmax, softmax nos indica que tanta probabilidad tiene cada clase, y por ende sabe
```

```
for i, layer in enumerate(model.layers):
    print(i, layer.name)
```

```
0 conv2d_1
1 max_pooling2d_1
2 conv2d_2
3 max_pooling2d_2
4 flatten_1
5 dense_1
6 dropout_1
7 dense_2
```

```
model.summary()
```

Model: "sequential_1"

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 160, 160, 32)	896
max_pooling2d_1 (MaxPooling2D)	(None, 80, 80, 32)	0
conv2d_2 (Conv2D)	(None, 80, 80, 64)	8256
max_pooling2d_2 (MaxPooling2D)	(None, 40, 40, 64)	0
flatten_1 (Flatten)	(None, 102400)	0

dense_1 (Dense)	(None, 256)	26214656
dropout_1 (Dropout)	(None, 256)	0
dense_2 (Dense)	(None, 7)	1799
=====		
Total params: 26,225,607		
Trainable params: 26,225,607		
Non-trainable params: 0		

```
data_entrenamiento = './data/entrenamiento'
data_validacion = './data/validacion'
```

```
train_datagen=ImageDataGenerator(preprocessing_function=preprocess_input) #included in our dependencies

train_generator=train_datagen.flow_from_directory(data_entrenamiento, # this is where you specify the path to the main data folder
                                                target_size=(160,160),
                                                color_mode='rgb',
                                                batch_size=16,
                                                class_mode='categorical',
                                                shuffle=True)

validacion_datagen = ImageDataGenerator(preprocessing_function=preprocess_input)

valid_generator = validacion_datagen.flow_from_directory(
    data_validacion,
    target_size=(160,160),
    batch_size=16,
    class_mode='categorical',
    shuffle=True)

print(train_generator.class_indices)
```

```
Found 1296 images belonging to 7 classes.
Found 324 images belonging to 7 classes.
{'Nivel_0': 0, 'Nivel_1': 1, 'Nivel_2': 2, 'Nivel_3': 3, 'Nivel_4': 4, 'Nivel_5': 5, 'Nivel_6': 6}
```

```
TOTAL_EPOCHAS_ENTRENAMIENTO = 16
```

```
model.compile(optimizer='Adam',loss='categorical_crossentropy',metrics=['accuracy'])
```

```
step_size_train=train_generator.n//train_generator.batch_size
step_size_valid=valid_generator.n//valid_generator.batch_size
```

```
H=model.fit_generator(generator=train_generator,
                    steps_per_epoch=step_size_train,
                    epochs=TOTAL_EPOCHAS_ENTRENAMIENTO,
                    validation_data=valid_generator,
                    validation_steps=step_size_valid)
```

```
Epoch 1/16
81/81 [=====] - 821s 10s/step - loss: 1.4267 - accuracy: 0.6188 - val_loss: 0.2413 - val_accuracy: 0.8469
Epoch 2/16
81/81 [=====] - 819s 10s/step - loss: 0.4439 - accuracy: 0.8279 - val_loss: 0.4052 - val_accuracy: 0.8117
Epoch 3/16
81/81 [=====] - 783s 10s/step - loss: 0.2630 - accuracy: 0.8981 - val_loss: 0.2934 - val_accuracy: 0.8864
Epoch 4/16
81/81 [=====] - 798s 10s/step - loss: 0.2329 - accuracy: 0.9159 - val_loss: 0.1415 - val_accuracy: 0.8961
Epoch 5/16
81/81 [=====] - 792s 10s/step - loss: 0.1520 - accuracy: 0.9452 - val_loss: 0.1688 - val_accuracy: 0.8929
Epoch 6/16
81/81 [=====] - 807s 10s/step - loss: 0.0968 - accuracy: 0.9660 - val_loss: 0.2638 - val_accuracy: 0.9026
Epoch 7/16
81/81 [=====] - 831s 10s/step - loss: 0.0665 - accuracy: 0.9738 - val_loss: 0.5100 - val_accuracy: 0.9416
Epoch 8/16
81/81 [=====] - 791s 10s/step - loss: 0.0388 - accuracy: 0.9931 - val_loss: 0.2060 - val_accuracy: 0.9318
Epoch 9/16
81/81 [=====] - 767s 9s/step - loss: 0.0300 - accuracy: 0.9884 - val_loss: 0.0507 - val_accuracy: 0.9123
Epoch 10/16
81/81 [=====] - 750s 9s/step - loss: 0.0879 - accuracy: 0.9630 - val_loss: 0.3540 - val_accuracy: 0.8409
Epoch 11/16
81/81 [=====] - 768s 9s/step - loss: 0.1040 - accuracy: 0.9622 - val_loss: 0.6276 - val_accuracy: 0.9221
Epoch 12/16
81/81 [=====] - 749s 9s/step - loss: 0.0828 - accuracy: 0.9684 - val_loss: 0.5933 - val_accuracy: 0.9253
```

```

Epoch 13/16
81/81 [=====] - 761s 9s/step - loss: 0.2078 - accuracy: 0.9406 - val_loss: 0.0104 - val_accuracy: 0.9091
Epoch 14/16
81/81 [=====] - 768s 9s/step - loss: 0.0659 - accuracy: 0.9784 - val_loss: 0.9818 - val_accuracy: 0.8799
Epoch 15/16
81/81 [=====] - 780s 10s/step - loss: 0.0492 - accuracy: 0.9823 - val_loss: 1.1041 - val_accuracy: 0.9026
Epoch 16/16
81/81 [=====] - 749s 9s/step - loss: 0.0768 - accuracy: 0.9676 - val_loss: 0.0019 - val_accuracy: 0.9286

```

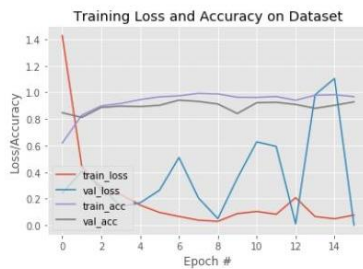
```
model.evaluate_generator(train_generator,train_generator.n//train_generator.batch_size)
```

```
[0.006610956974327564, 0.9961419701576233]
```

```

N = TOTAL_EPOCAS_ENTRENAMIENTO
plt.style.use("ggplot")
plt.figure()
plt.plot(np.arange(0, N), H.history["loss"], label="train_loss")
plt.plot(np.arange(0, N), H.history["val_loss"], label="val_loss")
plt.plot(np.arange(0, N), H.history["accuracy"], label="train_acc")
plt.plot(np.arange(0, N), H.history["val_accuracy"], label="val_acc")
plt.title("Training Loss and Accuracy on Dataset")
plt.xlabel("Epoch #")
plt.ylabel("Loss/Accuracy")
plt.legend(loc="lower left")
plt.savefig("simple/cnn_estadisticas_de_la_red.png")

```



==> SECCION DE PRUEBAS

```
test_datagen = ImageDataGenerator(preprocessing_function=preprocess_input)
```

```

test_generator = test_datagen.flow_from_directory(
    directory="./data/pruebas_expo/",
    target_size=(160, 160),
    color_mode="rgb",
    batch_size=1,
    class_mode=None,
    shuffle=False,
    seed=42
)

```

```
Found 35 images belonging to 1 classes.
```

```
STEP_SIZE_TEST=test_generator.n//test_generator.batch_size
```

```
test_generator.reset()
```

```
pred=model.predict_generator(test_generator, steps=STEP_SIZE_TEST, verbose=1) #devuelve el conjunto de prediccion
```

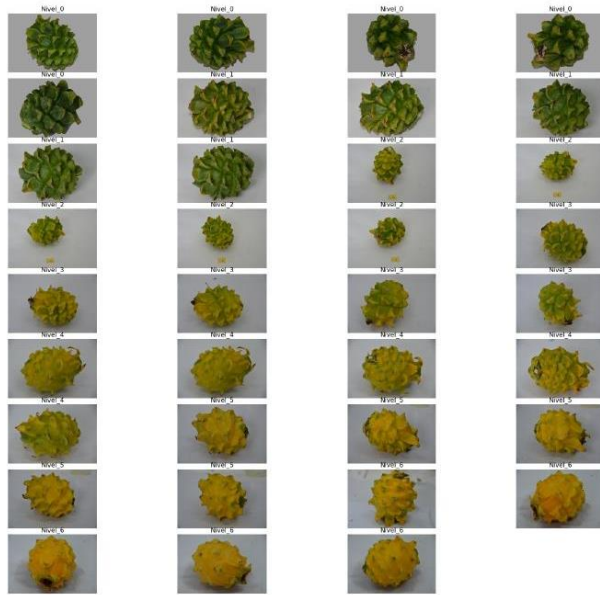
```
35/35 [=====] - 29s 828ms/step
```

```
pred.round(2)
```

```

array([[1. , 0. , 0. , 0. , 0. , 0. , 0. , 0. ],
       [1. , 0. , 0. , 0. , 0. , 0. , 0. , 0. ],
       [1. , 0. , 0. , 0. , 0. , 0. , 0. , 0. ],
       [1. , 0. , 0. , 0. , 0. , 0. , 0. , 0. ],
       [1. , 0. , 0. , 0. , 0. , 0. , 0. , 0. ],
       [0.01, 0.99, 0. , 0. , 0. , 0. , 0. , 0. ],
       [0. , 1. , 0. , 0. , 0. , 0. , 0. , 0. ],
       [0. , 1. , 0. , 0. , 0. , 0. , 0. , 0. ],
       [0. , 1. , 0. , 0. , 0. , 0. , 0. , 0. ]])

```

```

real_class_indices=[]
for i in range(0, len(filenamees)):
    #print ("++++",i,"++++")
    your_path = filenamees[i]
    path_list = your_path.split(os.sep)
    my_dict=train_generator.class_indices
    key_list = list(my_dict.keys())
    val_list = list(my_dict.values())
    #print(key_list[val_list.index(0)])
    #print (path_list[1])
    if (key_list[val_list.index(0)] in path_list[1]) :
        real_class_indices.append(0)
    if (key_list[val_list.index(1)] in path_list[1]) :
        real_class_indices.append(1)
    if (key_list[val_list.index(2)] in path_list[1]) :
        real_class_indices.append(2)
    if (key_list[val_list.index(3)] in path_list[1]) :
        real_class_indices.append(3)
    if (key_list[val_list.index(4)] in path_list[1]) :
        real_class_indices.append(4)
    if (key_list[val_list.index(5)] in path_list[1]) :
        real_class_indices.append(5)
    if (key_list[val_list.index(6)] in path_list[1]) :
        real_class_indices.append(6)

```

```

print (real_class_indices)
print (len(real_class_indices))
real_class_indices = np.array(real_class_indices)
print(type(real_class_indices))
[0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 6, 6, 6, 6]
35
<class 'numpy.ndarray'>

```

```
len(filenamees)
```

```
35
```

```
real_class_indices
```

```
array([0, 0, 0, 0, 0, 1, 1, 1, 1, 1])
```

```
len(predicted_class_indices)
```

```
35
```

```
cm = confusion_matrix(real_class_indices, predicted_class_indices)
```

```

def plot_confusion_matrix(cm, classes,
                          normalize=False,
                          title='Confusion matrix',
                          cmap=plt.cm.Blues):
    """
    Esta función imprime y traza la matriz de confusión.
    La normalización se puede aplicar configurando normalize=True.
    """
    plt.imshow(cm, interpolation='nearest', cmap=cmap)
    plt.title(title)
    plt.colorbar()
    tick_marks = np.arange(len(classes))
    plt.xticks(tick_marks, classes, rotation=45)
    plt.yticks(tick_marks, classes)

    if normalize:
        cm = cm.astype('float') / cm.sum(axis=1)[:, np.newaxis]
        print("Matriz de confusión normalizada")
    else:
        print("Matriz de confusión, sin normalización")

    print(cm)

    thresh = cm.max() / 2.
    for i, j in itertools.product(range(cm.shape[0]), range(cm.shape[1])):
        plt.text(j, i, cm[i, j],
                 horizontalalignment="center",
                 color="white" if cm[i, j] > thresh else "black")

    plt.tight_layout()
    plt.ylabel('Etiqueta verdadera')
    plt.xlabel('Etiqueta predecida')

```

```
cm_plot_labels = train_generator.class_indices
```

```
print (cm_plot_labels)
```

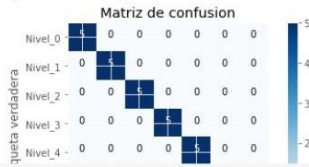
```
key_list = list(cm_plot_labels.keys())
```

```
val_list = list(cm_plot_labels.values())
```

```
{'Nivel_0': 0, 'Nivel_1': 1, 'Nivel_2': 2, 'Nivel_3': 3, 'Nivel_4': 4, 'Nivel_5': 5, 'Nivel_6': 6}
```

```
plot_confusion_matrix(cm, cm_plot_labels, title= 'Matriz de confusion')
```

```
Confusion matrix, without normalization
[[5 0 0 0 0 0]
 [0 5 0 0 0 0]
 [0 0 5 0 0 0]
 [0 0 0 5 0 0]
 [0 0 0 0 5 0]
 [0 0 0 0 0 5]]
```



```
print(real_class_indices)
print(predicted_class_indices)
```

```
[0 0 0 0 1 1 1 1 1 2 2 2 2 2 2 3 3 3 3 3 4 4 4 4 4 5 5 5 5 5 6 6 6 6 6]
[0 0 0 0 1 1 1 1 1 2 2 2 2 2 2 3 3 3 3 3 4 4 4 4 4 5 5 5 5 5 6 6 6 6 6]
```

etiqueta predicha

```
from mlxtend.evaluate import scoring
```

```
res = scoring(y_target=real_class_indices, y_predicted=predicted_class_indices, metric='error')
```

```
print('Error:', np.round(res * 100, 2), "%")
```

```
res = scoring(y_target=real_class_indices, y_predicted=predicted_class_indices, metric='accuracy')
print("Exactitud:", np.round(res*100, 2), "%")
```

```
res = scoring(y_target=real_class_indices, y_predicted=predicted_class_indices, metric='average per-class accuracy')
print("Promedio de exactitud por clase:", np.round(res*100, 2), "%")
```

```
Error: 0.0 %
Exactitud: 100.0 %
Promedio de exactitud por clase: 100.0 %
```

Autoguardado | cnn_pruebas_modelo.xlsx - Modo de compatibilidad | Buscar (Alt+Q)

Archivo Inicio Insertar Disposición de página Fórmulas Datos Revisar Vista Programador Ayuda

Portapapeles Fuente Alineación Número Estilos

	A	B	C	D	E	F	G	H	I	J	K
1	Filename	Predictions	Grados Brix	Acidez							
2	test_images\Nivel_0 (1).JPG	Nivel_0	3.6 - 12.19	1.5 - 2.01							
3	test_images\Nivel_0 (2).JPG	Nivel_0	3.6 - 12.19	1.5 - 2.01							
4	test_images\Nivel_0 (3).JPG	Nivel_0	3.6 - 12.19	1.5 - 2.01							
5	test_images\Nivel_0 (4).JPG	Nivel_0	3.6 - 12.19	1.5 - 2.01							
6	test_images\Nivel_0 (5).JPG	Nivel_0	3.6 - 12.19	1.5 - 2.01							
7	test_images\Nivel_1 (1).JPG	Nivel_1	13.9 - 15.8	0.19 - 0.24							
8	test_images\Nivel_1 (2).JPG	Nivel_1	13.9 - 15.8	0.19 - 0.24							
9	test_images\Nivel_1 (3).JPG	Nivel_1	13.9 - 15.8	0.19 - 0.24							
10	test_images\Nivel_1 (4).JPG	Nivel_1	13.9 - 15.8	0.19 - 0.24							
11	test_images\Nivel_1 (5).JPG	Nivel_1	13.9 - 15.8	0.19 - 0.24							
12	test_images\Nivel_2 (1).JPG	Nivel_2	13.8 - 16.4	0.12 - 0.23							
13	test_images\Nivel_2 (2).JPG	Nivel_2	13.8 - 16.4	0.12 - 0.23							
14	test_images\Nivel_2 (3).JPG	Nivel_2	13.8 - 16.4	0.12 - 0.23							
15	test_images\Nivel_2 (4).JPG	Nivel_2	13.8 - 16.4	0.12 - 0.23							
16	test_images\Nivel_2 (5).JPG	Nivel_2	13.8 - 16.4	0.12 - 0.23							
17	test_images\Nivel_3 (1).JPG	Nivel_3	14.2 - 17.5	0.11 - 0.22							
18	test_images\Nivel_3 (2).JPG	Nivel_3	14.2 - 17.5	0.11 - 0.22							
19	test_images\Nivel_3 (3).JPG	Nivel_3	14.2 - 17.5	0.11 - 0.22							
20	test_images\Nivel_3 (4).JPG	Nivel_3	14.2 - 17.5	0.11 - 0.22							
21	test_images\Nivel_3 (5).JPG	Nivel_3	14.2 - 17.5	0.11 - 0.22							
22	test_images\Nivel_4 (1).JPG	Nivel_4	11.2 - 17.6	0.09 - 0.16							
23	test_images\Nivel_4 (2).JPG	Nivel_4	11.2 - 17.6	0.09 - 0.16							
24	test_images\Nivel_4 (3).JPG	Nivel_4	11.2 - 17.6	0.09 - 0.16							
25	test_images\Nivel_4 (4).JPG	Nivel_4	11.2 - 17.6	0.09 - 0.16							
26	test_images\Nivel_4 (5).JPG	Nivel_4	11.2 - 17.6	0.09 - 0.16							
27	test_images\Nivel_5 (1).JPG	Nivel_5	11.5 - 17	0.1 - 0.17							
28	test_images\Nivel_5 (2).JPG	Nivel_5	11.5 - 17	0.1 - 0.17							
29	test_images\Nivel_5 (3).JPG	Nivel_5	11.5 - 17	0.1 - 0.17							
30	test_images\Nivel_5 (4).JPG	Nivel_5	11.5 - 17	0.1 - 0.17							
31	test_images\Nivel_5 (5).JPG	Nivel_5	11.5 - 17	0.1 - 0.17							
32	test_images\Nivel_6 (1).JPG	Nivel_6	14.8 - 16.7	0.1 - 0.15							
33	test_images\Nivel_6 (2).JPG	Nivel_6	14.8 - 16.7	0.1 - 0.15							
34	test_images\Nivel_6 (3).JPG	Nivel_6	14.8 - 16.7	0.1 - 0.15							
35	test_images\Nivel_6 (4).JPG	Nivel_6	14.8 - 16.7	0.1 - 0.15							
36	test_images\Nivel_6 (5).JPG	Nivel_6	14.8 - 16.7	0.1 - 0.15							

Sheet1

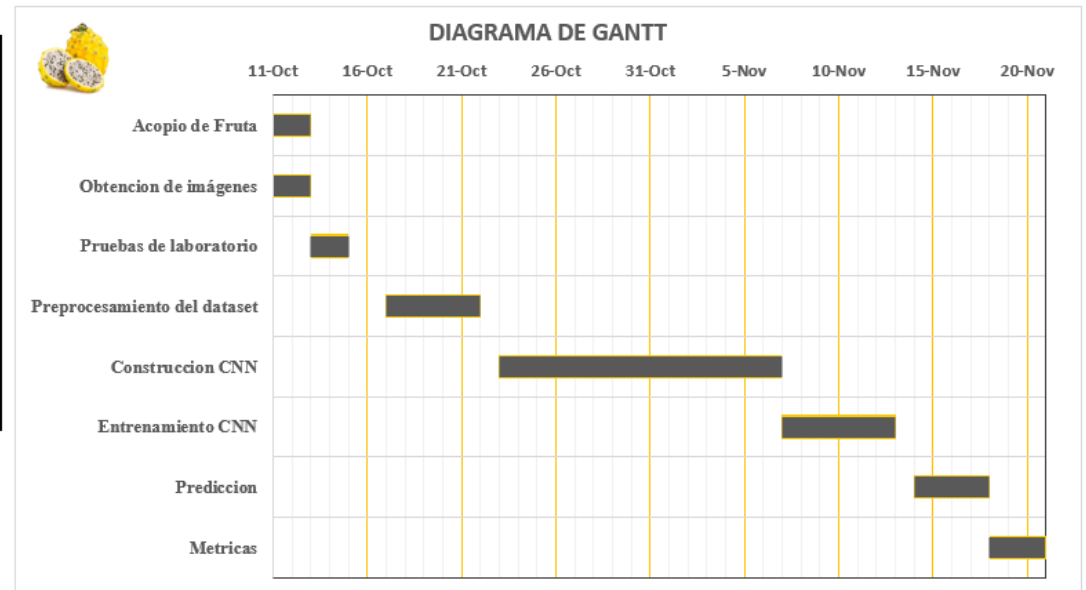
Anexo 5. Cronograma



CRONOGRAMA DE ACTIVIDADES DEL PROYECTO - REDES NEURONALES CONVOLUCIONALES PARA DETERMINAR LA ACIDEZ Y GRADOS BRIX DE PITAHAYA AMARILLA EN LA REGION AMAZONAS

FASE	ACTIVIDADES	FECHA INICIO	DURACION DIAS	FECHA FIN	% COMPLETADO	DIAS COMPLETADOS
Sprint 1	Acopio de Fruta	11-Oct	2	13-Oct	100%	2.00
Sprint 2	Obtencion de imágenes	11-Oct	2	13-Oct	100%	2.00
	Pruebas de laboratorio	13-Oct	2	15-Oct	100%	2.00
Sprint 3	Preprocesamiento del dataset	17-Oct	5	22-Oct	100%	5.00
	Construccion CNN	23-Oct	15	7-Nov	100%	15.00
	Entrenamiento CNN	7-Nov	6	13-Nov	100%	6.00
	Prediccion	14-Nov	4	18-Nov	85%	4.00
	Metricas	18-Nov	3	21-Nov	90%	3.00

INICIO PROYECTO	44845.00
FIN PROYECTO	44886.00





UNIVERSIDAD CÉSAR VALLEJO

**FACULTAD DE INGENIERÍA Y ARQUITECTURA
ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS**

Declaratoria de Autenticidad del Asesor

Yo, NECOCHEA CHAMORRO JORGE ISAAC, docente de la FACULTAD DE INGENIERÍA Y ARQUITECTURA de la escuela profesional de INGENIERÍA DE SISTEMAS de la UNIVERSIDAD CÉSAR VALLEJO SAC - LIMA NORTE, asesor de Tesis titulada: "Redes Neuronales Convolucionales para determinar la Acidez y Grados Brix de Pitahaya Amarilla (*Hylocereus megalanthus*) en la Región Amazonas", cuyo autor es CUEVA CARO JEISSON ENRIQUE, constato que la investigación tiene un índice de similitud de 10.00%, verificable en el reporte de originalidad del programa Turnitin, el cual ha sido realizado sin filtros, ni exclusiones.

He revisado dicho reporte y concluyo que cada una de las coincidencias detectadas no constituyen plagio. A mi leal saber y entender la Tesis cumple con todas las normas para el uso de citas y referencias establecidas por la Universidad César Vallejo.

En tal sentido, asumo la responsabilidad que corresponda ante cualquier falsedad, ocultamiento u omisión tanto de los documentos como de información aportada, por lo cual me someto a lo dispuesto en las normas académicas vigentes de la Universidad César Vallejo.

LIMA, 29 de Noviembre del 2022

Apellidos y Nombres del Asesor:	Firma
NECOCHEA CHAMORRO JORGE ISAAC DNI: 18167347 ORCID: 0000-0002-3290-8975	Firmado electrónicamente por: JNECOCHEA el 29- 11-2022 16:58:55

Código documento Trilce: TRI - 0461468