



CARTA DE AUTORIZACIÓN

CÓDIGO

AP-BIB-FO-06

VERSIÓN

1

VIGENCIA

2014

PÁGINA

1 de 2

Neiva, _____ 25/10/2021 _____

Señores

CENTRO DE INFORMACIÓN Y DOCUMENTACIÓN

UNIVERSIDAD SURCOLOMBIANA

Ciudad

El (Los) suscrito(s):

_____ Cesar Mauricio Gutierrez Diaz _____, con C.C. No. _____ 1075305868 _____,

_____ María Camila Villegas Gómez _____, con C.C. No. _____ 1077875293 _____,

_____, con C.C. No. _____,

_____, con C.C. No. _____,

Autor(es) de la tesis y/o trabajo de grado o _____ Tesis _____

titulado__ DESARROLLO DE UN MODELO DE DISCRIMINACIÓN DE DEFECTOS DE CAFÉ, A TRAVÉS DEL ENTENDIMIENTO DE INFORMACIÓN FÍSICOQUÍMICA Y DE IMÁGENES EN EL ESPECTRO VISUAL _____

presentado y aprobado en el año ____ 2021 _____ como requisito para optar al título de

_____ Ingeniero Agrícola _____;

Autorizo (amos) al CENTRO DE INFORMACIÓN Y DOCUMENTACIÓN de la Universidad Surcolombiana para que, con fines académicos, muestre al país y el exterior la producción intelectual de la Universidad Surcolombiana, a través de la visibilidad de su contenido de la siguiente manera:

- Los usuarios puedan consultar el contenido de este trabajo de grado en los sitios web que administra la Universidad, en bases de datos, repositorio digital, catálogos y en otros sitios web, redes y sistemas de información nacionales e internacionales "open access" y en las redes de información con las cuales tenga convenio la Institución.
- Permita la consulta, la reproducción y préstamo a los usuarios interesados en el contenido de este trabajo, para todos los usos que tengan finalidad académica, ya sea en formato Cd-Rom o digital desde internet, intranet, etc., y en general para cualquier formato conocido o por conocer, dentro de los términos establecidos en la Ley 23 de 1982, Ley 44 de 1993, Decisión Andina 351 de 1993, Decreto 460 de 1995 y demás normas generales sobre la materia.

Vigilada Mineducación



CARTA DE AUTORIZACIÓN

CÓDIGO

AP-BIB-FO-06

VERSIÓN

1

VIGENCIA

2014

PÁGINA

2 de 2

- Continúo conservando los correspondientes derechos sin modificación o restricción alguna; puesto que, de acuerdo con la legislación colombiana aplicable, el presente es un acuerdo jurídico que en ningún caso conlleva la enajenación del derecho de autor y sus conexos.

De conformidad con lo establecido en el artículo 30 de la Ley 23 de 1982 y el artículo 11 de la Decisión Andina 351 de 1993, "Los derechos morales sobre el trabajo son propiedad de los autores", los cuales son irrenunciables, imprescriptibles, inembargables e inalienables.

EL AUTOR/ESTUDIANTE:

Cesar Mauricio Gutierrez Diaz

Firma: 

EL AUTOR/ESTUDIANTE:

María Camila Villegas Gómez

Firma: 

EL AUTOR/ESTUDIANTE:

Firma: _____

EL AUTOR/ESTUDIANTE:

Firma: _____



DESCRIPCIÓN DE LA TESIS Y/O TRABAJOS DE GRADO

CÓDIGO

AP-BIB-FO-07

VERSIÓN

1

VIGENCIA

2014

PÁGINA

1 de 3

TÍTULO COMPLETO DEL TRABAJO: DESARROLLO DE UN MODELO DE DISCRIMINACIÓN DE DEFECTOS DE CAFÉ, A TRAVÉS DEL ENTENDIMIENTO DE INFORMACIÓN FÍSICOQUÍMICA Y DE IMÁGENES EN EL ESPECTRO VISUAL

AUTOR O AUTORES:

Primero y Segundo Apellido	Primero y Segundo Nombre
Gutierrez Diaz	Cesar Mauricio
Villegas Gómez	María Camila

DIRECTOR Y CODIRECTOR TESIS:

Primero y Segundo Apellido	Primero y Segundo Nombre
Girón Hernández	Joel
Sarria González	Juan Camilo

ASESOR (ES):

Primero y Segundo Apellido	Primero y Segundo Nombre

PARA OPTAR AL TÍTULO DE: Ingeniero Agrícola

FACULTAD: Ingeniería

PROGRAMA O POSGRADO: Agrícola

CIUDAD: Neiva (Huila)

AÑO DE PRESENTACIÓN: 2021 **NÚMERO DE PÁGINAS:**22

TIPO DE ILUSTRACIONES (Marcar con una X):

Diagramas___Fotografías___Grabaciones en discos_x___Ilustraciones en general_x___Grabados___
Láminas___Litografías___Mapas___Música impresa___Planos___Retratos___ Sin ilustraciones_x___Tablas
x o Cuadros_x_

Vigilada Mineducación

La versión vigente y controlada de este documento, solo podrá ser consultada a través del sitio web Institucional www.usco.edu.co, link Sistema Gestión de Calidad. La copia o impresión diferente a la publicada, será considerada como documento no controlado y su uso indebido no es de responsabilidad de la Universidad Surcolombiana.



DESCRIPCIÓN DE LA TESIS Y/O TRABAJOS DE GRADO

CÓDIGO	AP-BIB-FO-07	VERSIÓN	1	VIGENCIA	2014	PÁGINA	2 de 3
---------------	---------------------	----------------	----------	-----------------	-------------	---------------	---------------

SOFTWARE requerido y/o especializado para la lectura del documento: No

MATERIAL ANEXO: No

PREMIO O DISTINCIÓN (En caso de ser LAUREADAS o Meritoria): Meritoria

PALABRAS CLAVES EN ESPAÑOL E INGLÉS:

<u>Español</u>	<u>Inglés</u>	<u>Español</u>	<u>Inglés</u>
1. café de calidad	coffee quality	6. _____	_____
2. defectos de café	green beans	7. _____	_____
3. procesamiento de imagen	image processing	8. _____	_____
4. PCA	PCA	9. _____	_____
5. Redes Neuronales	neural networks	10. _____	_____

RESUMEN DEL CONTENIDO: (Máximo 250 palabras)

Eliminar los granos defectuosos en una carga de café verde aumentan la probabilidad de obtener resultados sobresalientes en su respectivo análisis sensorial. La clasificación de los granos generalmente se realiza manualmente mediante inspección visual o utilizando seleccionadoras mecánicas. Estos procedimientos presentan ciertas limitaciones como: subjetividad, confiabilidad intermedia, tiempos de trabajo prolongados e incremento de costos. El objetivo de este trabajo consistió en: i) obtener información fisicoquímica y de imagen en el espectro visual de las muestras de café defectuoso y no defectuoso, ii) evaluar químico métricamente mediante la información físico-química y digital de las muestras, iii) desarrollar un modelo de discriminación mediante la construcción de una red neuronal convolucional (CNN) utilizándola información del espectro visual obtenido de los granos. Los resultados mostraron que el PCA permitió comprimir la información arrojada por las imágenes explicando el 82.3 % de la variabilidad de las muestras con los nueve primeros componentes, únicamente diferenciando los granos no defectuosos de las categorías de defectos negro y cardenillo. La precisión general de la CNN fue de 92 %, los granos control se detectaron con una precisión del 95% y una recuperación del 100%.

ABSTRACT: (Máximo 250 palabras)

Eliminating defective beans in a batch of green coffee increases the likelihood of outstanding results in the respective sensory analysis. Grading is usually done manually by visual inspection or using mechanical selectors. These procedures have certain limitations such as: subjectivity, intermediate reliability, time consuming and high-costs. The objective of this work consisted of: i) obtaining physicochemical and image information in the visual spectrum of defective and non-defective beans, ii) evaluating the information obtained from the samples by principal component analysis (PCA), iii) developing a discrimination model through the



DESCRIPCIÓN DE LA TESIS Y/O TRABAJOS DE GRADO

CÓDIGO	AP-BIB-FO-07	VERSIÓN	1	VIGENCIA	2014	PÁGINA	3 de 3
---------------	---------------------	----------------	----------	-----------------	-------------	---------------	---------------

construction of a convolutional neural network (CNN) using information from the visual spectrum obtained from the beans. The results showed that the PCA was able to compress the information provided by the images, explaining 82.3 % of the variability of the samples with the first nine components, only differentiating the non-defective beans from the black and cardenillo defect categories. The overall accuracy of the CNN was 92 %, the control kernels were detected with an accuracy of 95 % and a recovery of 100 %.

APROBACIÓN DE LA TESIS

Firma: *Diana Carolina Polania Montiel*

Nombre presidente del Jurado: Diana Carolina Polania Montiel

Firma:

Nombre Jurado: Yelson Fernando Barrios Rodríguez

DESARROLLO DE UN MODELO DE DISCRIMINACIÓN DE DEFECTOS DE CAFÉ, A TRAVÉS DEL ENTENDIMIENTO DE INFORMACIÓN FÍSICOQUÍMICA Y DE IMÁGENES EN EL ESPECTRO VISUAL

Gutiérrez-Díaz, C. M¹, Villegas-Gómez, M. C², Girón-Hernández, J³, Sarria-González, J.C⁴

Resumen

Eliminar los granos defectuosos en una carga de café verde aumentan la probabilidad de obtener resultados sobresalientes en su respectivo análisis sensorial. La clasificación de los granos generalmente se realiza manualmente mediante inspección visual o utilizando seleccionadoras mecánicas. Estos procedimientos presentan ciertas limitaciones como: subjetividad, confiabilidad intermedia, tiempos de trabajo prolongados e incremento de costos. El objetivo de este trabajo consistió en: i) obtener información fisicoquímica y de imagen en el espectro visual de las muestras de café defectuoso y no defectuoso, ii) evaluar químicamente mediante la información físico-química y digital de las muestras, iii) desarrollar un modelo de discriminación mediante la construcción de una red neuronal convolucional (CNN) utilizándola información del espectro visual obtenido de los granos. Los resultados mostraron que el PCA permitió comprimir la información arrojada por las imágenes explicando el 82.3 % de la variabilidad de las muestras con los nueve primeros componentes, únicamente diferenciando los granos no defectuosos de las categorías de defectos negro y cardenillo. La precisión general de la CNN fue de 92 %, los granos control se detectaron con una precisión del 95% y una recuperación del 100%.

Palabras clave: café de calidad, defectos de café, procesamiento de imagen, PCA, redes neuronales.

Abstract

Eliminating defective beans in a batch of green coffee increases the likelihood of outstanding results in the respective sensory analysis. Grading is usually done manually by visual inspection or using mechanical selectors. These procedures have certain limitations such as: subjectivity, intermediate reliability, time consuming and high-costs. The objective of this work consisted of: i) obtaining physicochemical and image information in the visual spectrum of defective and non-defective beans, ii) evaluating the information obtained from the samples by principal component analysis (PCA), iii) developing a discrimination model through the construction of a convolutional neural network (CNN) using information from the visual spectrum obtained from the beans. The results showed that the PCA was able to compress the information provided by the images, explaining 82.3 % of the variability of the samples with the first nine components, only differentiating the non-defective beans from the black and cardenillo defect categories. The overall accuracy of the CNN was 92 %, the control kernels were detected with an accuracy of 95 % and a recovery of 100 %.

Keywords: coffee quality, green beans, image processing, PCA, neural networks.

Introducción

El café es el producto agrícola más importante en los mercados mundiales (Giacalone et al, 2019), en general, una bebida de café de alta calidad está compuesta esencialmente por granos sin defectos (Yi-fang, 2012). El café excelso en Colombia se define como granos de color verde y con tamaño similar, con una tolerancia de hasta 5% de almendras retenidas entre las mallas 14/64” y 12/64”; esto quiere decir que son granos de dimensiones de 5,6 y 4,76 mm que pasan por la criba 14 y 12. Por ello es necesario la clasificación de los granos por su tamaño y daños, como lo establece la resolución 05 del 2002 de la Federación Nacional de Cafeteros de Colombia (FNC), los criterios de tamaño del grano y el número máximo admisible de defectos en el café a exportar. Por su parte la asociación de cafés especiales (SCA por sus siglas en inglés) establece la guía de defectos para café verde Arábica lavado, incluyendo dos categorías de defectos. La categoría I, que castiga severamente la presencia de defectos, comprende los siguientes tipos de granos: negro, agrio (avinagrado), cereza seca, daño por hongo, materia extraña y dañado gravemente por insectos (brocado). Entre los defectos más comunes de la categoría 2 se encuentran los granos partidos, que se generan

durante el proceso de despulpado o el proceso de molienda seca, donde el ajuste del equipo es incorrecto y se aplica fricción o presión excesivas partiendo los granos.

Los defectos pueden también causar alteraciones químicas dentro del grano de café teniendo relación con la calidad química del café, esta se asocia principalmente al contenido de ácidos clorogénicos, cafeína y trigonelina (Garg et al 2021; Pimpley et al., 2020; Wianowska & Gil 2019). Estos compuestos, se encuentran en altas concentraciones en los granos, asociándose a propiedades organolépticas, estimulantes y a efectos terapéuticos de la bebida (Costa et al., 2020). Adicionalmente, la acidez titulable ha demostrado una correlación lineal entre los sólidos disueltos totales con la percepción de acidez en café extraído por goteo (Batali et al., 2021). Se ha demostrado que los granos defectuosos presentan niveles más altos de sacarosa, trigonelina y ácidos clorogénicos, lo que genera un perfil de sabor más amargo y astringente (Geromel et al., 2006). Por tanto, eliminar los granos defectuosos que se puedan presentar en una carga de café, es imprescindible para obtener una bebida sensorialmente satisfactoria, y así lograr un precio de venta favorable.

La selección de los granos defectuosos usualmente se realiza en cooperativas, asociaciones o industrias procesadoras de café, mediante métodos manuales y mecánicos (Godinez, 2017). La selección manual se realiza por personal especializado, lo que requiere una inversión considerable de tiempo y dinero (Gomes, 2014). Adicionalmente, es un trabajo intensivo que llega a ser arduo, laborioso y subjetivo. Se ha comprobado que una clasificación realizada dos veces por un mismo operario presenta una alta probabilidad del 85% de ser diferente (Paulus et al., 1997). Por su parte, la clasificación mecánica se realiza con equipos que diferencian los granos por tamaño y densidad mediante vibraciones. Esta selección es poco precisa, no evalúan parámetros de forma ni de color (Baigvand et al., 2015) y puede causar daño mecánico a los granos evaluados debido al contacto directo con la máquina (Amiriparian et al., 2008).

Por tanto, se hace necesario desarrollar métodos para clasificar los granos defectuosos de los no defectuosos, de manera simple, rápida, económica y confiable. Sánchez – Aguilar et al., (2019) destacan la necesidad de velar por las características del producto que llega al

consumidor final, enfatizando en la automatización de los diferentes procesos de control de calidad. En este sentido, la visión por computador ha sido empleada en el desarrollo de aplicaciones que permiten optimizar procesos en la agricultura y la industria alimentaria en las etapas de cultivo, cosecha y postcosecha (Kakani, et al, 2020).

La visión por computador se basa en el desarrollo de algoritmos que extraen y analizan automáticamente la información presente en una imagen y la comparan con un patrón definido (Khan, et al, 2018). En general, un algoritmo de visión artificial está conformado por dos etapas principales: i) preprocesamiento de las imágenes, empezando por una segmentación precisa de los objetos de interés del fondo/ruido y ii) extracción de la información que caracteriza el objeto, (García et al, 2019).

En la industria alimentaria el análisis de imagen permite obtener un alto volumen de información a bajo costo (Zhou et al., 2019). El procesamiento de la información puede realizarse con métodos no supervisados como el análisis de componentes principales (PCA) o métodos de aprendizaje supervisado, entre los que se encuentran los modelos conocidos como redes neuronales. Estos últimos métodos de clasificación son similares a las conexiones neurológicas del cerebro humano (Acevedo et al., 2017). El modelo se compone de neuronas (unidad computacional) que se organizan en grupos (capas), que interactúan con neuronas de otras capas conformando una red (Matich, 2001). La información de entrada en una red se transmite por las neuronas de las capas mediante operaciones matemáticas lineales y no lineales hasta producir la clasificación de los datos a la salida.

Las redes neuronales convolucionales (CNN) han demostrado ser útiles en el análisis de imagen, debido a que organizan los datos de entrada en forma de matrices, como puede ser imágenes en blanco y negro o en el sistema de color rojo (R), verde (G) y azul (B) (RGB), permitiendo la extracción de características de color y forma para realizar tareas de clasificación (Sikakollu & Dash 2021).

En la clasificación de imágenes de granos de café mediante redes neuronales, se han realizado trabajos con diferentes enfoques: i) diferenciación entre granos defectuosos y no defectuosos

ii) diferentes categorías de granos defectuosos y, iii) determinando categorías de calidad de los granos según características de tamaño, forma y color. Esto se logró desarrollando patrones de comparación (MLP) con base en características de color y forma, logrando una precisión del 94.1% (Pizzaia et al., 2018). En esta línea, se propuso un sistema automático de selección en tiempo real utilizando una CNN logrando una precisión del 93% (Huang et al., 2019). Así mismo se han entrenado CNN logrando precisiones de clasificación entre 72.4% y 98.7% para granos defectuosos (Pinto et al., 2017). Otros investigadores compararon diferentes herramientas como support vector machine, deep neural network y random forest, presentando precisiones mayores al 88% para los clasificadores evaluados (Dos Santos et al., 2020). Por ejemplo, se han entrenado redes neuronales artificiales (ANN por sus siglas en inglés) que lograron niveles de clasificación entre 40% y 100% (Faridah et al., 2011). También se ha realizado la evaluación de las arquitecturas VGG16 y ResNet 152, para evaluar la construcción de una CNN, obteniendo la mejor precisión para ResNet-152 (67.8%). Otros autores acoplaron ANN con un clasificador de Bayes, lo que les permitió clasificar granos de café verde por color con precisiones del 100%. Sin embargo, a pesar de que se ha demostrado la eficacia de algunos modelos para diferenciar entre diferentes tipos de granos de café o defectos, hasta el momento, los estudios realizados se presentan directamente desde la construcción de sistemas de visión artificial sin tener en cuenta la caracterización fisicoquímica de las muestras y sin realizar análisis multivariante de la información de las imágenes. Por tanto, en este trabajo se contempló: i) seleccionar las muestras, realizar la adquisición de imágenes y caracterización fisicoquímica, ii) evaluar químicamente mediante la información físico química y digital obtenida para diferenciar entre granos defectuosos y no defectuosos, y iii) construir un sistema de diferenciación y clasificación de defectos mediante CNN.

1. Materiales y métodos

2.1 Selección y preparación de las muestras

Se utilizaron granos de café *Coffea arabica* sin defectos (control) y defectos de café de diferentes variedades, provenientes de la cosecha de 2019 de diferentes municipios del Huila. Se seleccionaron manualmente 428 granos control (78,84 g) y granos defectuosos definidos según la SCA (2018): 1140 granos negros (182 g), 281 avinagrados (44,96 g), 105 cereza

seca (14,7g), 258 cardenillos (43,86 g), 825 brocados (140,25 g); y 656 mordidos/partidos (104,96 g).

2.2 Obtención y procesamiento de las imágenes de las muestras

Adquisición de las imágenes

La figura 1 muestra el esquema del montaje utilizado para la obtención de las imágenes. Se empleó un dispositivo de carga acoplado (CCD), que es un circuito integrado con un número delimitado de condensadores acoplados, que puede transferir su carga a otros condensadores internamente en el circuito, compuesto de 1,5 a 2,5 millones de píxeles de 20 a 70 μm ; el cual tiene los píxeles que se encuentran en forma de matriz generando una capa de centelleo que incurren en los fotones de rayos X y se convierten posteriormente en Luz, esta luz interactúa con el silicio creando un paquete de carga, estos personalizan la imagen latente y se transfieren en un amplificador de lectura a un convertidor análogo – digital del ordenador (Trevejo, 2016). Canon T3 (Tokio, Japón) y un sistema con lámparas LED de iluminación fría Philips Corepro LED 12.5 W con temperatura de color de 6500°K (Eindhoven, Países Bajos). Para la adquisición de imágenes, la lente se calibró con una apertura de $f/3$, a una distancia focal de aproximadamente 41 mm y una profundidad de campo de alrededor 3 cm, las capturas se realizaron de manera remota, para evitar descalibrar la precisión del sistema mediante el software Capture One PRO (Phase one, Dinamarca)

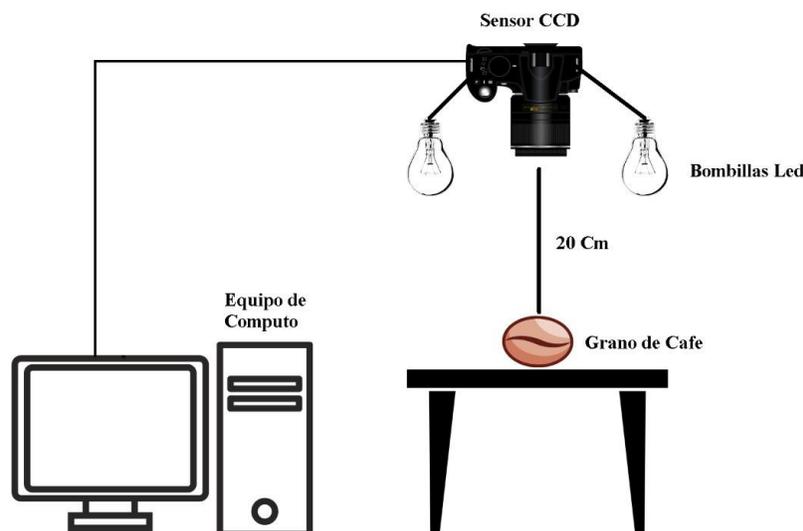


Figura 1. Esquema de la configuración del sistema de captura de imagen.

Procesamiento de la información digital

Lectura dinámica del directorio de imágenes. Se desarrolló un algoritmo para procesar las imágenes adquiridas utilizando el software MATLAB 8.0 (Mathworks, EE. UU.). Se creó un vector denominado “*vectornombres*” para almacenar los datos de las capturas, desarrollando un bucle que buscaba, leía y analizaba de manera independiente cada una de las imágenes almacenadas en dicho vector.

Segmentación y recorte de las imágenes. La figura 2 muestra la segmentación dinámica de las imágenes para diferenciar el objeto de interés del fondo. El proceso consistió en transformar las coordenadas de color acrónimo Red Green Blue (RGB) a escala de grises (Forsyth & Ponce, 2003), se aplicó un umbral dinámico de Otsu para binarizar cada imagen (Otsu, 1979); posteriormente, se eliminó el relleno aplicando una máscara que segmentaba a la imagen capturada. Finalmente, se etiquetó cada uno de los granos y se recortó en una nueva imagen con un tamaño de 250x250 píxeles de 50 μm para ingresar la información al sistema de clasificación.

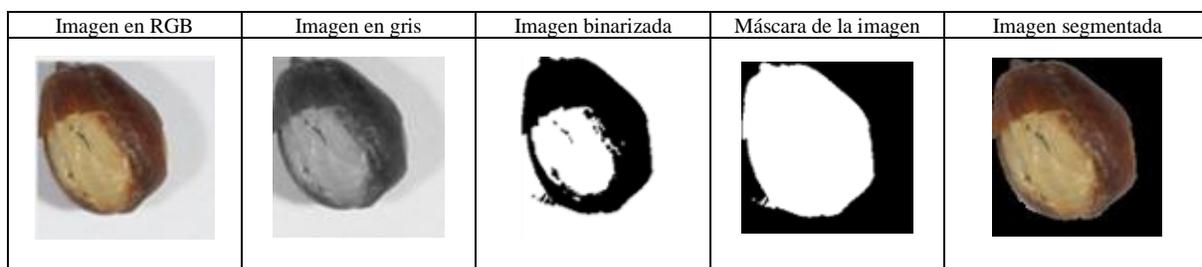


Figura 2. Proceso de segmentación de la imagen para el análisis de la información.

Obtención de la información de cada imagen. Las imágenes se procesaron con un algoritmo adicional que extrajo un vector de datos con información de forma y de color. Se obtuvieron 50 valores diferentes, 30 variables para cuantificar los cambios en el color en los valores mínimo, medios, máximos, medianos y desviaciones de las coordenadas rojo, verde, azul (RGB) y la caracterización espectral de un escáner en la adquisición de coordenadas de

la imagen de un objeto cromático dando los atributos valor, saturación y tono por sus siglas en inglés (VHS); facilitando la interpretación del escáner y otros dispositivos de despliegue e impresión de imágenes (Bermúdez et al., 2010). Adicionalmente, se emplearon métodos de morfología matemática obteniendo seis características físicas de los granos (área, área convexa, excentricidad, perímetro, solidez, equidiámetro) y 14 variables de textura de Haralick (momento angular segundo/Homogeneidad, contraste, momento diferencial inverso, entropía, correlación, varianza en Haralick, suma de promedios, suma de entropía, diferencia de entropía, diferencia de varianza, suma de varianza, información de medidas de correlación, agrupación de sombras, cluster prominence) (Andrews, 1971). Los vectores obtenidos para el procesamiento de cada grano se recopilaban en una matriz única de datos, obteniendo la información de imagen para el análisis de componentes principales.

2.3 Caracterización fisicoquímica de las muestras

En la caracterización fisicoquímica de las muestras se utilizaron aproximadamente 45 gramos de cada una de las categorías evaluadas. Las muestras se trituraron en un molino de café G3 HD (Bunn, EE.UU.), con un tamaño de partícula de 200 micrones, esta medida se logró mediante un molido en fino durante 5 minutos. Las determinaciones se realizaron por triplicado.

Determinación de color. Las coordenadas de color L^*a^*b del espacio CIELab, se obtuvieron de manera directa con un colorímetro CR-400 (Konica Minolta, Japón), con iluminante estándar D45. Según lo estipulado en la NTC 2442.

Determinación de pH, acidez titulable y sólidos solubles. La cuantificación del pH, sólidos solubles y la acidez titulable se realizó mediante un extracto con 5 g de cada muestra en 100 ml de agua desionizada caliente (90 °C), siguiendo lo establecido por la NTC 5247. Los sólidos solubles se determinaron de manera directa poniendo 1 ml del extracto en un refractómetro digital PR-201 alfa (Atago, Japón), los valores obtenidos se expresaron en grados Brix (°Bx). Para la determinación del pH y acidez titulable, la solución (extracto) se agitó durante 3 minutos procurando homogeneizar. Finalmente, se filtró y se llevó a un volumen de 100 ml en un balón aforado. La medición de pH se realizó cuando la disolución

alcanzó temperatura ambiente ($25\pm 2^{\circ}\text{C}$) con un medidor de pH digital starter 300C (Ohaus, EE.UU.). La acidez titulable se determinó mediante adición de hidróxido de sodio (0.1N) hasta que la solución alcanzó un valor de pH 6.5, calculada mediante la ecuación 1, expresado en mg de ácido clorogénico/g de café.

$$\text{Acidez titulable} = \frac{\text{Volumen de NaOH} * \text{Equivalente ácido}}{\text{Volumen de muestra}} \quad (\text{Ecuación 1})$$

Determinación de ácidos clorogénicos (ACGs), cafeína y trigonelina. La cuantificación de los ACGs, cafeína y trigonelina se realizó mediante cromatografía de líquida de alta resolución (HPLC); fue necesario un extracto de 2 g de cada muestra en 20 ml de agua bidestilada caliente (90°C), se agito en un baño ultrasonido (modelo S 30 H Elmasonic, Alemania) durante 20 minutos; las muestras se filtraron utilizando un filtro de celulosa regenerada de $0.45\ \mu\text{m}$ (Fuentes et al., 2018). Se utilizó un cromatógrafo 1260 Infinity II LC System (Agilent, EEUU.) equipado con un auto-muestreador (modelo 1260 vilsampler) y un detector de UV (modelo 1260 DAD HS). La separación se realizó durante 24 minutos con una columna C18 ($150 \times 4.6\ \text{mm i.d. } 5\ \mu\text{m}$) (Agilent, EEUU.). La fase móvil utilizada consistió en un método de elución isocrático (25:75, v/v) MeOH y agua bidestilada acidificada con ácido acético glacial (0.1 %). Se realizaron lecturas a una longitud de onda de 324 nm para ACGs, la cafeína y trigonelina a 268nm, en ambos casos se inyectaron $20\ \mu\text{L}$ con un flujo de $1.2\ \text{mL min}^{-1}$. Para la cuantificación de los tres analitos se realizaron curvas de calibrado de 5 puntos en rangos de concentración de 20 a 500 ppm, con un R^2 promedio de 0.999.

2.4 Análisis estadístico

La información obtenida de las determinaciones fisicoquímicas se estudió mediante ANOVA simple, para observar si existían diferencias entre las categorías establecidas de las muestras, con un nivel de confianza del 95 %. Se incluyó un análisis post-hoc de mínima diferencia significativa de Fisher (LSD), para establecer los valores medios diferentes en los tratamientos. Adicionalmente, la información obtenida de las determinaciones fisicoquímicas y la información del análisis de imagen se evaluaron separadamente mediante el análisis de componentes principales (PCA). Con la información fisicoquímica, se estudió la influencia

y la asociación de las variables sobre la diferenciación de las categorías en análisis. Mientras que la información digital se analizó con el propósito de observar las agrupaciones obtenidas. Los análisis se realizaron con el software estadístico Statgraphics centurión XVI (Manugistics Inc., EE.UU.).

2.4.1 Construcción del sistema

Se construyó una red CNN con arquitectura VGG16 entrenada con un subconjunto de datos iniciales, para esto se empleó una estrategia de ajuste fino que clasificó las muestras en siete categorías, granos defectuosos (6 categorías) y no defectuosos. La primera capa de la red presentó una buena relación entre eficiencia y complejidad. El ajuste minimizó el tiempo del entrenamiento debido a que aprovechó las características aprendidas del grupo inicial de datos. Posteriormente, se realizó el procedimiento de entrenamiento de la red, el primer paso consistió en un calentamiento que permitió actualizar los pesos de la capa inicial, después del entrenamiento/inicialización en esta capa, se continuó con el entrenamiento del resto de la red.

Finalmente, se empleó un optimizador de descenso de gradiente estocástico con una tasa de aprendizaje de 0.0004 y un impulso de 0.9. El valor de la pérdida se calculó como una activación de Softmax más una pérdida de CrossEntropy. Lo que dio como resultado la CNN entrenada, que generó una probabilidad sobre las categorías para cada imagen. Este tipo de pérdida se usa comúnmente para la clasificación de clases múltiples. El marco de aprendizaje profundo empleado fue Keras 2.2.4 con un backend TensorFlow GPU versión 1.14 en un sistema operativo Microsoft Windows 10.

2. Resultados y discusión

3.1 Determinaciones analíticas

La tabla 1 muestra los valores promedio obtenidos de las determinaciones químicas realizadas a las muestras. En los resultados de sólidos solubles, el control presentó valores similares a los de las categorías de granos partidos, avinagrados y cereza seca. Los menor valores se presentaron para los defectos de granos negro y brocados. Los resultados obtenidos

para los granos negros se pueden atribuir a una menor concentración de azúcares y carbohidratos en este tipo de defecto, como lo ha reportado Vasconcelos et al., 2007 quienes encontraron un 92 y 87.5 % menos de sacarosa en los granos negros comparados con granos no defectuosos y avinagrados. En los granos brocados el insecto (*H. hampei*) entra en los frutos del café perforando el tejido del pericarpio en el endospermo, disminuyendo sus compuestos (Góngora et al., 2020), lo que lleva como consecuencia a una disminución de los sólidos solubles. En los resultados obtenidos para la acidez, los valores se encuentran dentro de los rangos reportados (Clifford, 1985), el control presentó los menores valores en comparación con el resto de las categorías. Freitas, 2001 atribuyó la baja acidez titulable de los granos verdes en buenas condiciones a la ausencia de mucílago en este tipo de grano, por lo que no hay fermentación de azúcares de esta capa, que originen ácidos que puedan difundirse en la semilla, en cuanto a la baja acidez de los granos negros, comparada con los otros defectos podría ser atribuido a la pérdida (lixiviación o volatilización) de ácidos durante el contacto de este grano con el suelo (Vasconcelos et al. 2007). Según estudios previos, la acidez debería aumentar a medida que la calidad del café disminuye (Franca et al., 2005; Mazzafera, 1999).

Ramalakshmi, 2007 menciona que la alta acidez de los granos con defecto cereza seca se atribuye a la presencia prolongada de mucílago en este tipo de grano, ya que estos pasan por un proceso de secado en planta o en el suelo manteniendo su mucilago, por lo que hay fermentación excesiva de azúcares de esta capa, que originan ácidos que se difunden a través de la semilla. Por el contrario, el valor de pH encontrado para los granos control fue mayor que el resto de las categorías y concuerda con los rangos reportados en otros estudios 5.26-6.10 (Bicho et al., 2013); adicionalmente, se considera que valores altos de pH están asociados con mejor calidad del café, lo que concuerda con los resultados del control frente a las muestras defectuosas. Esta disminución de pH en los defectos en comparación con los valores de los granos control están asociadas a procesos del grado de maduración, la especie y otros factores asociados a la calidad del café, tal como la altura y la presencia o ausencia de sombra e inclusive se les relaciona con la resistencia a algunas enfermedades y el estrés que estas causan al cafeto (Solis, 2002).

Tabla 1. Valores medios y desviaciones estándar obtenidos de las determinaciones de sólidos solubles, acidez, pH, ácido clorogénico, cafeína y trigonelina.

Muestra	Sólidos solubles	Acidez (mg ACGs/g café)	pH	ACGs (mg/g café)	Cafeína (mg/g café)	Trigonelina (mg/g café)
Control	0.86±0.05 ^{cd}	4.25±0.01 ^a	6.03±0.03 ^f	39.75±2.06 ^b	9.61±0.39 ^b	10.25±0.27 ^a
Negro	0.53±0.06 ^a	12.75±1.42 ^d	5.30±0.05 ^b	53.03±4.71 ^{cd}	11.92±0.96 ^c	10.94±0.81 ^{ab}
Avinagrado	0.90±0.03 ^d	13.92±1.74 ^d	5.54±0.06 ^c	55.91±3.08 ^e	12.56±0.44 ^c	13.66±0.44 ^{cd}
Cereza seca	0.82±0.02 ^c	20.29±1.64 ^e	5.13±0.12 ^a	31.31±7.62 ^a	6.87±2.57 ^a	11.81±2.57 ^{abc}
Cardenillo	1.25±0.07 ^e	13.22±0.82 ^d	5.57±0.05 ^c	51.78±4.09 ^{cd}	13.05±0.85 ^c	13.93±0.79 ^d
Broca	0.70±0.03 ^b	9.92±1.42 ^c	5.74±0.03 ^d	45.96±1.70 ^{bc}	11.58±0.20 ^c	11.91±0.09 ^{abcd}
Partido	0.84±0.02 ^{cd}	6.58±0.80 ^b	5.87±0.03 ^e	45.42±6.71 ^{bc}	12.42±1.25 ^c	12.88±1.12 ^{bcd}

Letras diferentes dentro de la misma columna indican la existencia de diferencias significativas ($P < 0.05$).

El contenido de ácidos clorogénicos y cafeína de los granos control (Tabla 1), fue similar a lo encontrado en otras investigaciones, solamente la cereza seca presentó menores valores, las demás categorías fueron significativamente más altas, la variación en el contenido de ácidos clorogénicos depende de la especie evaluada, del grado de maduración de frutos y, en menor medida de las prácticas de siembra y cosecha, clima y suelo (Clifford, 1985; Farah et al., 2005), por otra parte Iriondo et al., 2018 nos dice que el aumento en el contenido de ACGs y cafeína en los granos defectuosos se debe a una respuesta fisiológica defensiva de la planta. Finalmente, la trigonelina fue el compuesto en el que se observaron menores diferencias significativas entre categorías; sin embargo, el valor de los granos control fue ligeramente menor al de las categorías avinagrado, cardenillo y partido, estos datos concuerdan con Farah et al. (2006) quien tampoco encontró diferencias significativas en los niveles de trigonelina entre cafés de alta y baja calidad, esta mínima variación entre los diferentes defectos y el café control en el valor de la trigonelina están ligados a la especie y su concentración está fuertemente influenciada por la variedad genética (Avelino et al., 2005; Leroy et al., 2006).

Los resultados de la evaluación de las variables físicas de las muestras se observan en la tabla 2. La coordenada L^* presentó diferencias para cada una de las categorías evaluadas, el control se posicionó únicamente por encima de los granos defectuosos negros, este comportamiento se debe a que el resto de las categorías suele presentar colores más claros. Los valores de la

coordenada a*, fueron similares para el control, los brocados y los partidos, una agrupación similar se observó para la coordenada b*.

Tabla 2. Media más la desviación estándar de las coordenadas de color L*a*b* (CIELab).

Muestra	L	a*	b*
Control	45.85±0.21 ^b	1.90±0.02 ^b	18.58±0.63 ^a
Negro	43.99±0.04 ^a	6.17±0.01 ^d	18.66±0.03 ^a
Avinagrado	60.92±0.76 ^e	-2.57±1.37 ^a	32.33±2.39 ^c
Cereza seca	50.58±0.11 ^c	7.80±0.05 ^e	24.02±0.22 ^b
Cardenillo	57.73±0.11 ^d	4.09±0.03 ^c	23.49±0.09 ^b
Broca	60.83±0.11 ^e	1.70±0.01 ^b	19.64±0.02 ^a
Partido	63.35±0.09 ^f	1.59±0.02 ^b	18.44±0.04 ^a

Letras diferentes dentro de la misma columna indican la existencia de diferencias significativas (P<0.05).

Los resultados del análisis de color no permitieron diferenciar el control de categorías como brocado y los partidos. En una carga de café es importante retirar los granos brocados, este defecto es considerado individualmente en el cálculo del factor de rendimiento y genera una penalización (Hena, 2008). por otro lado Franca y col. (2005) informaron que los valores de luminosidad eran más altos para los frijoles molidos en comparación con los enteros (medidas de superficie), lo que indica que la superficie del frijol es más oscura que su interior, y que las diferencias de luminosidad entre superficie e interior eran más significativas para los granos defectuosos en comparación con los no defectuosos, esto explica los altos niveles de luminancia que tiene los defectos como el negro comparados con los datos del control. Por tanto, la información ofrecida por este conjunto de variables no es suficiente para realizar la clasificación de las muestras, siendo uno de los principales desafíos de las máquinas seleccionadoras ópticas que actualmente se utilizan para la clasificación de granos. Por tanto, se necesitan controles que tengan en cuenta otros parámetros.

3.2 Análisis de componentes principales de las variables fisicoquímicas

El PCA realizado con los resultados de las determinaciones fisicoquímicas, explicó el 82.3 % de la variabilidad de las muestras con los primeros tres componentes principales (CP). La figura 3 exhibe la distribución de las muestras. El CP1 (40.38 %) permitió diferenciar los defectos negros asociados al valor de la coordenada de color a*, mientras el CP2 (27.96 %) agrupó el control, los granos brocados y los granos partidos, estas dos últimas categorías se asociaron a la variable pH. La información permite inferir que los defectos avinagrados y

cardenillos se localizan en el cuadrante I, asociándose con ACGs, trigonellina, sólidos solubles y las coordenadas de color b^* y L^* . En el cuadrante II se agruparon los defectos de cereza seca relacionados con la variable de acidez titulable. El resultado de este análisis permite observar que las muestras control no se asociaron a ninguna de las variables evaluadas. Por tanto, para el desarrollo de un sistema de clasificación por visión artificial, se debe ir más allá de la clasificación por coordenadas de color dado que esta variable se asocia a diferentes categorías de los defectos sin permitir diferenciar efectivamente los granos utilizados como controles.

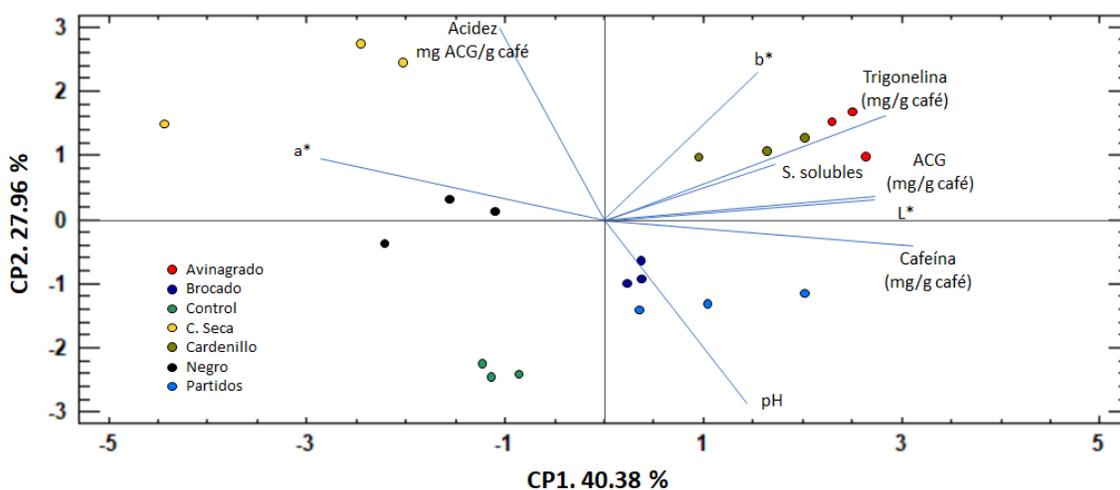


Figura 3. Diagrama de dispersión con los componentes y las muestras del PCA obtenido de las variables fisicoquímicas.

3.3 Procesamiento de la información digital.

En el análisis de la información de las fotografías, se eliminaron las características constantes para todas las categorías de las muestras con el propósito de evitar la colinealidad. En este procedimiento se excluyeron las variables de color de: máximos y mínimos de las coordenadas tonalidad (H) y saturación (S), los máximos de la coordenada B (sistema RGB), las variables de forma: solidez, equidiámetro, las variables texturales Halarick de: correlación, varianza en Haralick, suma de entropía, diferencia de entropía, información de medidas de correlación, agrupación de sombras, reduciendo así la información de 50 a 37 variables.

3.4 Análisis de principales de la información digital

La figura 4 muestra el PCA resultante de las imágenes adquiridas, con dos componentes principales se explicó el 46.5% del comportamiento del modelo. El incremento de componentes por encima de este número no mejora la varianza explicada. El CP1 (34.2 %) permitió una separación de los defectos de las categorías cardenillo y negros. Los controles se ubicaron por debajo del CP2 (12.3%); sin embargo, no se evidenció la separación de los controles con las categorías de granos avinagrados, partidos, brocados y cereza seca, lo que dificulta la discriminación de las muestras. Por tanto, se hace necesario establecer un clasificador que permita diferenciar claramente los granos control del resto.

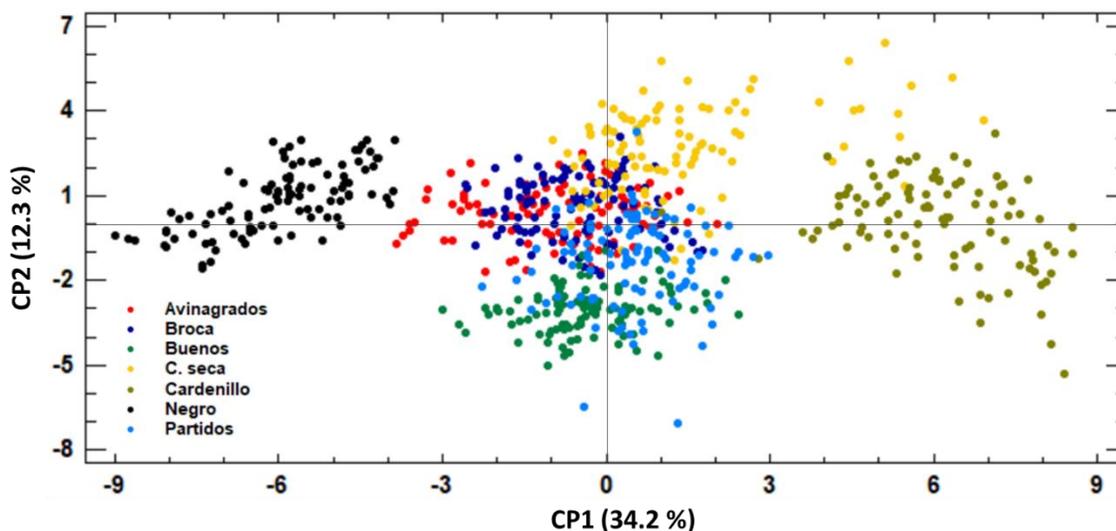


Figura 4. Análisis de componentes principales de la información obtenida de las imágenes de las muestras de café.

3.5 Entrenamiento y evaluación de la red convolucional (CNN)

La figura 5 muestra los resultados del entrenamiento (A) y validación (B). Durante la fase de entrenamiento, la precisión aumentó de forma continua y casi lineal hasta el 94% con la iteración 10. A partir de ese momento el aumento de precisión se ralentizó hasta que se mantuvo en valores constantes de 96% a partir de la iteración 18.

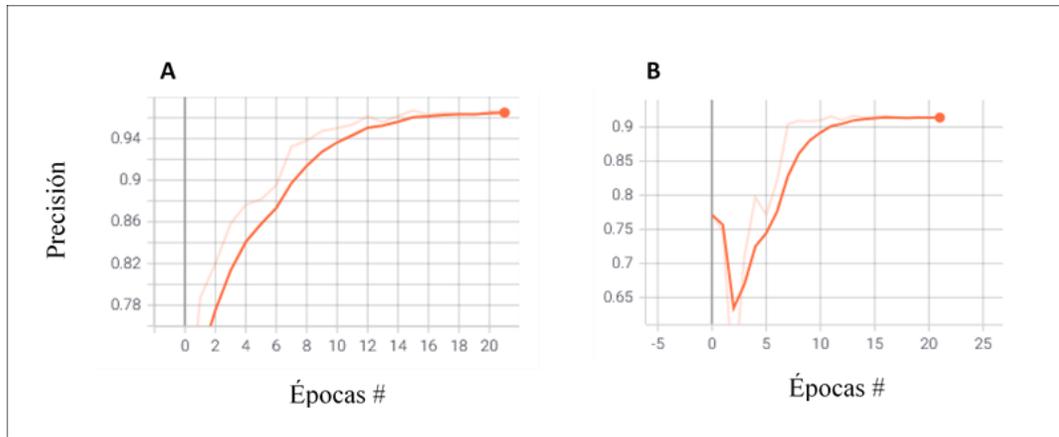


Figura 5. Resultados del entrenamiento (A) y validación (B) de la red neuronal convolucional construida para la clasificación de granos control y defectos de café.

En el grupo de validación, que permitió controlar el proceso de aprendizaje, disminuyó drásticamente por debajo del 65% de precisión aproximadamente a las 5 iteraciones. Sin embargo, después de 10 épocas se aproximó al 90% de precisión, manteniéndose constantemente con muy poca fluctuación. La disminución drástica en la validación se debió a que la red neural aprendió características muy específicas de las muestras de entrenamiento que no aparecían en las muestras de validación. Estas características específicas se suprimieron por otras más genéricas al seguir entrenando más iteraciones. Se puede afirmar que el sistema es válido porque en el entrenamiento no se produce este fenómeno y en la validación la precisión se mantiene constante a partir de la iteración 10.

La tabla 3 muestra los resultados de la clasificación obtenida con la red entrenada, la precisión general de la red fue de 92% y el peso de cada uno de los componentes estuvo entre 90 y 92 %. Los mejores resultados se observaron para las categorías de cereza seca y los granos control. En ambos casos la precisión fue mayor igual al 95 % y la capacidad del clasificador para detectar los granos que cumplían con las características determinadas para cada categoría fue de 100 %. Adicionalmente, los valores del F1-score para las categorías en mención fueron las mejores de la red, lo significa que la recuperación y la precisión tuvieron la misma importancia.

Tabla 3. Métricas de clasificación obtenidas en la validación de la red neuronal entrenada para la selección de granos defectuosos de café

Categoría	Precisión	Recuperación	F1-score
Brocados	0.93	0.91	0.92
Control	0.95	1.00	0.97
Cardenillo	0.98	0.83	0.90
Cereza seca	1.00	1.00	1.00
Partidos	0.87	0.84	0.85
Negros	0.94	0.99	0.96
Avinagrados	0.89	0.90	0.90

La clasificación de los granos partidos presentó los menores rendimientos en el grupo de muestras evaluadas, sin embargo, las métricas de clasificación estuvieron por encima del 84%. La red evaluada cumple con el objetivo de detectar y seleccionar los granos control, esto es importante debido a que en el campo práctico lo que se busca es la eliminación de cualquier clase de defecto que se pueda presentar.

4 Conclusiones

La metodología utilizada propuesta para la identificación de granos defectuosos utilizando visión artificial con el entrenamiento de la red neuronal propuesta (CNN), permitió diferenciar los granos control, mostrando una precisión del fue de 92 %. Se demostró que tanto las fisicoquímicas con el análisis no supervisado de las imágenes (PCA), presentan limitaciones para realizar una clasificación certera. La metodología propuesta mediante la CNN surge como alternativa a los métodos de selección utilizados actualmente, clasificadoras ópticas y mecánicas. Siendo una alternativa económica, rápida y objetiva, aplicable a la selección de granos de café.

- **5 Bibliografía** Acevedo, E., Serna, A., & Serna, E. (2017). Principios y características de las redes neuronales artificiales. *Desarrollo e innovación en ingeniería*, 173.
- Amiriparian, J., Khoshtaqaza, M.H., Kabir, E., & Minaee, S. (2008). Estimación de volumen de patatas con procesamiento de imágenes. *J. Agric. Eng. Res.* 8 (4), 113–126.
- Andrews, H. C. (1971). Multidimensional rotations in feature selection. *IEEE Transactions on Computers*, 100(9), 1045-1051.

- Araújo, C. D. S., Macedo, L. L., Vimercati, W. C., Ferreira, A., Prezotti, L. C., & Saraiva, S. H. (2020). Determination of pH and acidity in green coffee using near-infrared spectroscopy and multivariate regression. *Journal of the Science of Food and Agriculture*, 100(6), 2488-2493.
- Avelino J, Barboza B, Araya JC, Fonseca C, Davrieux F, Guyot B, Cilas C (2005) Effects of slope exposure, altitude and yield on coffee quality in two altitude terroirs of Costa Rica, Orosi and Santa María de Dota. *J Sci Food Agric*. 85:1869-1876.
- Baigvand, M., Banakar, A., Minaei, S., Khodaei, J., & Behrooz-Khazaei, N. (2015). Sistema de visión artificial para la clasificación de higos secos. *Comput. Electron. Agric.*, 119, 158–165
- Batali, M. E., Cotter, A. R., Frost, S. C., Ristenpart, W. D., & Guinard, J. X. (2021). Titratable Acidity, Perceived Sourness, and Liking of Acidity in Drip Brewed Coffee. *ACS Food Science & Technology*.
- Bicho, N. C., Leitão, A. E., Ramalho, J. C., de Alvarenga, N. B., & Lidon, F. C. (2013). Identification of chemical clusters discriminators of Arabica and Robusta green coffee. *International Journal of Food Properties*, 16(4), 895-904.
- Costa, M. C., Lima, T. F. O., Arcaro, C. A., Inacio, M. D., Batista-Duharte, A., Carlos, I. Z., ... & Baviera, A. M. (2020). Trigonelline and curcumin alone, but not in combination, counteract oxidative stress and inflammation and increase glycation product detoxification in the liver and kidney of mice with high-fat diet-induced obesity. *The Journal of nutritional biochemistry*, 76, 108303.
- Dos Santos, F., Rosas, J., Martins, R., Araújo, G., Viana, L., & Gonçalves, J. (2020). Quality assessment of coffee beans through computer vision and machine learning algorithms. *Coffee Science-ISSN 1984-3909*, 15, e151752-e151752.
- Farah, A., Monteiro, M.C., Calado, V., Franca, A.S. and Trugo, L.C. (2006). Correlation between cup quality and chemical attributes of Brazilian coffee. *Food chemistry*, 98, 373-380.
- Faridah, F., Parikesit, G. O., & Ferdiansjah, F. (2015 de 2011). Coffee bean grade determination based on image parameter. *TELKOMNIKA (Telecommunication*

Computing Electronics and Control), 9(3), 547-554.
doi:10.12928/telkomnika.v9i3.747

- Forsyth, D. A., & Ponce, J. (2003). A modern approach. *Computer vision: a modern approach*, 17, 21-48.
- Franca, A. S., Oliveira, L. S., Mendonça, J. C. F., & Silva, X. A. (2005). Physical and chemical attributes of defective crude and roasted coffee beans. *Food Chemistry*, 90(1-2), 84-89.
- Fuentes López, A., Fuentes López, C., & García Martínez, E. M. (2018). Determinación de cafeína en café mediante cromatografía líquida de alta resolución (HPLC).
- Gamboa Rosero, P. Y., Mosquera Sánchez, S. A., & Paz Narváez, I. E. (2015). Caracterización física de café especial (*Coffea Arabica*) en el municipio de Chachagüí (Nariño, Colombia). *Revista Lasallista de Investigación*, 12(1), 90-98.
- García, M., Candelo-Becerra, J. E., & Hoyos, F. E. (2019). Quality and defect inspection of green coffee beans using a computer vision system. *Applied Sciences*, 9(19), 4195.
- Garg, S. K., Shukla, A., & Choudhury, S. (2021). Green coffee beans. In *Nutraceuticals* (pp. 725-748). Academic Press.
- Geromel, C., Ferreira, L. P., Guerreiro, S. M. C., Cavalari, A. A., Pot, D., Pereira, L. F. P., ... & Marraccini, P. (2006). Biochemical and genomic analysis of sucrose metabolism during coffee (*Coffea arabica*) fruit development. *Journal of Experimental Botany*, 57(12), 3243-3258.
- Giacalone, D., Degn, T. K., Yang, N., Liu, C., Fisk, I., & Münchow, M. (2019). Common roasting defects in coffee: Aroma composition, sensory characterization and consumer perception. *Food quality and preference*, 71, 463-474.
- Godínez, K. I. (2017). Modelo de empresa procesadora y comercializadora de café especial nacional con particularidad de coffee shop.
- Gomes, L. (2014). " Efecto de diferentes técnicas de secado sobre la calidad del café arábica (*Coffea Arabica* L.) en Timor-Leste", Dissertação para obtenção do grau de mestre em Engenharia Agronomica. Universidade de Evora Escola de Ciência e Tecnologia Evora Portugal.

- Góngora, C. E., Tapias, J., Jaramillo, J., Medina, R., Gonzalez, S., Casanova, H., ... & Benavides, P. (2020). Evaluation of terpene-volatile compounds repellent to the coffee berry borer, *Hypothenemus hampei* (Ferrari)(Coleoptera: Curculionidae). *Journal of Chemical Ecology*, 46(9), 881-890.
- Henao, L. (2008). Control de la broca del café *Hypothenemus hampei* (Ferrari) con extractos vegetales de plantas de la flora regional(Tesis Tecnólogo en Química). Universidad Tecnológica de Pereira, Facultad de Tecnología: Pereira (Colombia), 70 pp.
- Huang KY, Wu CH, Hong QB, et al (2019) Speech Emotion Recognition Using Deep Neural Network Considering Verbal and Nonverbal Speech Sounds. In: ICASSP, IEEE.
- Iriondo-DeHond, A., Aparicio García, N., Fernandez-Gomez, B., Guisantes-Batan, E., Velázquez Escobar, F., Blanch, G. P., ... del Castillo, M. D. (2018). Validation of coffee by-products as novel food ingredients. *Innovative Food Science & Emerging Technologies*. doi:10.1016/j.ifset.2018.06.010
- Kakani, V., Nguyen, V. H., Kumar, B. P., Kim, H., & Pasupuleti, V. R. (2020). A critical review on computer vision and artificial intelligence in food industry. *Journal of Agriculture and Food Research*, 2, 100033.
- Khan, M. J., Khan, H. S., Yousaf, A., Khurshid, K., & Abbas, A. (2018). Modern trends in hyperspectral image analysis: A review. *IEEE Access*, 6, 14118-14129.
- Leroy T, Ribeyre F, Bertrand B, Charmetant P, Dufour M, Montagnon C, Marraccini P, Pot D (2006) Genetics of coffee quality. *Braz J Plant Physiol*. 18:229-242
- Matich, D. J. (2001). Redes Neuronales: Conceptos básicos y aplicaciones. *Universidad Tecnológica Nacional, México*, 41.
- Matich, D. J. (2001). Redes Neuronales: Conceptos básicos y aplicaciones. *Universidad Tecnológica Nacional, México*, 41.
- NTC 5247. (2004). Café tostado, en grano o molido. determinación de la acidez titulable. Instituto Colombiano de Normas Técnicas y Certificación (ICONTEC), Bogotá.
- Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1), 62-66.

- P Paulus, I., De Busscher, R., & Schrevens, E. (1997). Use of image analysis to investigate human quality classification of apples. *Journal of Agricultural Engineering Research*, 68(4), 341-353.
- Pimpley, V., Patil, S., Srinivasan, K., Desai, N., & Murthy, P. S. (2020). The chemistry of chlorogenic acid from green coffee and its role in attenuation of obesity and diabetes. *Preparative Biochemistry & Biotechnology*, 1–10. doi:10.1080/10826068.2020.1786699
- Pinto, C., Furukawa, J., Fukai, H., & Tamura, S. (2017, August). Classification of Green coffee bean images basec on defect types using convolucional neural network (CNN). In *2017 International Conference on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA)* (pp. 1-5). IEEE.
- Pizzaia, J. P. L., Salcides, I. R., de Almeida, G. M., Contarato, R., & de Almeida, R. (2018, November). Arabica coffee samples classification using a Multilayer Perceptron neural network. In *2018 13th IEEE International Conference on Industry Applications (INDUSCON)* (pp. 80-84). IEEE.
- Ramalakshmi, K., Kubra, I.R. and Rao, L.J.M. (2007). Physicochemical Characteristics of Green Coffee: Comparison of Graded and Defective Beans, *Journal of Food Science*, 72, S333-S337
- Sánchez-Aguiar, A. F., Ceballos-Arroyo, A. M., & Espinosa-Bedoya, A. (2019, April). Toward the recognition of non-defective coffee beans by means of digital image processing. In *2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA)* (pp. 1-5). IEEE.
- SCA. (2018). *El café Arábica lavado Guía de defectos del café verde*.
- Sikakollu, P., & Dash, R. (2021). Ensemble of multiple CNN classifiers for HSI classification with Superpixel Smoothing. *Computers & Geosciences*, 154, 104806.
- Solís, (2002). Desarrollo de un método de análisis para la cuantificación de ácidos clorogénicos en café. *Agronomía Costarricense* 29(2): 99-107.[Documento en línea].Disponible: http://www.mag.go.cr/rev_agr/v29n02_099.pdf.

- Vasconcelos, A. L. S., Franca, A. S., Glória, M. B. A., & Mendonça, J. C. (2007). A comparative study of chemical attributes and levels of amines in defective green and roasted coffee beans. *Food Chemistry*, 101(1), 26-32.
- Wianowska, D., & Gil, M. (2019). Recent advances in extraction and analysis procedures of natural chlorogenic acids. *Phytochemistry Reviews*, 18(1), 273-302
- Yi-Fang Chu. (2012). *Café: Efectos emergentes en la salud y prevención de enfermedades*, First Edition John Wiley & Sons, Blackwell Publishing Inc, Iowa.
- Zhou, L., Zhang, C., Liu, F., Qiu, Z., & He, Y. (2019). Application of deep learning in food: a review. *Comprehensive Reviews in Food Science and Food Safety*, 18(6), 1793-1811.