

AlgarroboNet - cuando menos es más

AlgarroboNet - When less is more

Resumen—La identificación de especies forestales es de creciente interés, para conservar especies en peligro de extinción como el Algarrobo (*Neltuma pallida*). Esto, representa una problemática al darle seguimiento a los árboles semilleros conocidos como árboles Plus. Para este fin, los vehículos aéreos no tripulados (VANT) han mostrado su utilidad, especialmente al acoplarse a técnicas como las redes neuronales convolucionales (RNC). Sin embargo, la creciente complejidad de las RNC conlleva a la necesidad de contrastar el rendimiento frente a su complejidad. El objetivo fue determinar si existe diferencias en las métricas de rendimiento del proceso de clasificación de árboles de algarrobo plus en RNC de creciente complejidad estructural. Para este fin se generó una base de datos de imágenes de árboles plus y no plus de algarrobo. Posteriormente, usando la base de datos se implementaron modelos de clasificación aplicando la transferencia de entrenamiento a dos RNC de diferente complejidad (Alexnet y Googlenet) y una mediante entrenamiento “since scratch” a una propuesta de RNC de complejidad intermedia denominada *AlgarroboNet*. El entrenamiento se realizó treinta veces, aleatorizando los datos de modelado y validación, calculando en todos los casos las métricas *accuracy* y $F2_{measure}$. El número de capas y parámetros entrenables de la RNC *GoogleNet* fue 8 y 4.2 veces superior al de *AlgarroboNet*; a pesar de ello, comparando sus tiempos de entrenamiento, esta solo requirió la mitad del tiempo. Asimismo, las métricas de entrenamiento de la red *GoogleNet* fueron muy ligeramente mejores, superando en 4 %, a las reportadas para *AlgarroboNet*. Los resultados demuestran que es necesario evaluar la complejidad de la tarea (número de clases, homogeneidad intra y extra clases, etc.) antes de elegir la RNC, a fin de evitar sobredimensionar el modelo.

Palabras clave - Algarrobo, clasificación de árboles, Red Neuronal Convolucional, Transferencia de entrenamiento

Abstract— *The identification of forest species is a problem of growing interest, mainly due to its application to the monitoring of seed trees (Plus) in endangered species such as the Carob (Neltuma pallida). For this end, Unmanned aerial vehicles (UAV) have shown their usefulness, especially when coupled with techniques such as convolutional neural networks (RNC). Therefore, The objective was to determine if there are differences in the performance metrics of the carob plus tree classification process in RNC of increasing structural complexity. For this purpose, a database of images of plus and non-plus carob trees was generated. Subsequently, using the database, classification models were implemented by applying the transfer of training to two RNCs of different complexity (Alexnet and GoogleNet) and one through "since scratch" training to a proposal for RNC of intermediate complexity called AlgarroboNet. The training was carried out thirty times, randomizing the modeling and validation data, calculating in all cases the accuracy and $F2_{measure}$ metrics. The number of layers and trainable parameters of RNC GoogleNet was 8 and 4.2 times higher than that of AlgarroboNet; Despite this, its training required only a half the time compared to AlgarroboNet. Likewise, the training metrics of the GoogleNet network were very slightly better, exceeding by 4 %, those reported for AlgarroboNet. The results show that it is necessary to evaluate the complexity of the task (number of classes,*

intra- and extra-class homogeneity, etc.) before choosing the RNC, in order to avoid oversizing the model.

Keywords - Carob tree, tree classification, Convolutional Neural Network, Training transfer

I. INTRODUCCIÓN

El objetivo 15 de la Agenda Sostenible 2030 busca proteger y restablecer los ecosistemas terrestres, mientras que el objetivo 6 prioriza la protección de los ecosistemas relacionados con el agua. Dos importantes metas que motivan esta investigación son: (1) velar por la conservación y restablecimiento de los ecosistemas terrestres, y (2) adoptar medidas urgentes y significativas para detener la degradación de los hábitats naturales, evitar la pérdida de la diversidad biológica, proteger a las especies en extinción, y restablecer los ecosistemas relacionados con el agua, incluidos los bosques [1].

En línea con estos objetivos, la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO, por sus siglas en inglés) promueve un creciente interés en los procesos de reforestación de áreas boscosas [2], [3]. Sin embargo, persisten los procesos de deforestación que afectan gravemente a especies en peligro de extinción como el algarrobo (*Neltuma pallida*), cuya importancia es crucial en ecosistemas frágiles como los bosques secos [4]. En este escenario, la reforestación requiere un inventario de las poblaciones de algarrobos, lo cual implica identificar los *plus árboles* como fuentes de semillas [5].

La identificación de árboles de algarrobo *plus* se basa en la medición *in-situ* de características como tamaño de copa, porcentaje de tallos sanos, diámetro y altura del tronco, entre otras [6]. Este proceso es demandante de tiempo y personal capacitado, razón por la cual la teledetección basada en imágenes obtenidas por satélites, aviones, y vehículos aéreos no tripulados (VANT) se ha mostrado especialmente útil [7], [8], [9].

Estas técnicas de análisis, implementadas mediante herramientas computacionales, suelen incluir aprendizaje máquina (*Machine Learning*), especialmente diseñadas para extraer información y organizarla como base en la toma de decisiones [10], [11], [12]. Como sugieren Benos [13] y Debus [14], el enfoque de aprendizaje profundo (*Deep Learning*), que utiliza redes neuronales convolucionales (RNC), es capaz de extraer información y auto-organizarse, posibilitando tareas como clasificación y predicción.

Un gran número de algoritmos utilizados en el análisis de imágenes pertenecen a la categoría de enfoques RNC, los cuales han demostrado ser muy eficaces en el sector forestal

para identificar y separar árboles de otra vegetación mediante imágenes capturadas por VANT [11], [15].

En general, los enfoques de entrenamiento de CNN implican la construcción de arquitecturas con varios parámetros, distribuidos en muchas capas de procesamiento. Los modelos resultantes suelen ser complejos en términos de recursos de memoria, con el consiguiente alto coste computacional.

En función de lo expuesto, se propuso en este estudio evaluar la relación entre la complejidad de la estructura y los parámetros y métricas de entrenamiento de RNC para su uso en la clasificación de árboles *plus* y *no plus* de algarrobo.

II. MATERIALES Y METODOLOGÍA

El procedimiento metodológico se desarrolló siguiendo la secuencia ilustrada en la Figura 1, implementando rutinas y funciones mediante scripts en Matlab 2023a (MathWorks, Inc.).

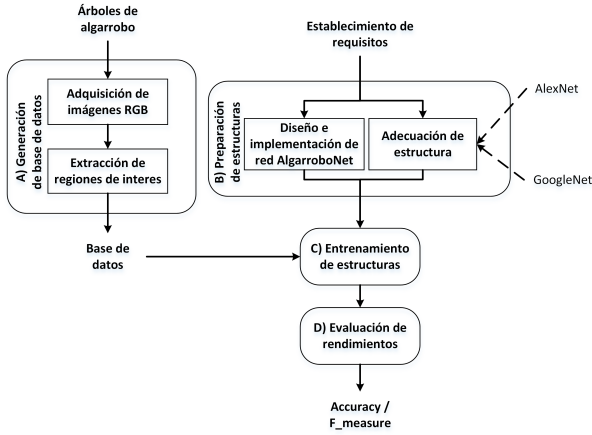


Figura 1: Secuencia de la metodología empleada

A continuación, se detallan cada una de las etapas del proceso experimental.

II-A. Generación de base de datos

La identificación y clasificación in situ de árboles de algarrobo *plus* y *no plus* se realizó en áreas bajo la gestión del Servicio Nacional Forestal (SERFOR), siguiendo las características fenotípicas propuestas por [16]. Las imágenes en *RGB* se adquirieron utilizando un VANT Phantom 4 RTK (DJI, China).

Posteriormente, las imágenes fueron corregidas para reducir el efecto del desplazamiento y segmentadas manualmente para identificar las copas de los árboles seleccionados, creando regiones de interés (ROI) de forma análoga a lo descrito por [17], [18]. Las ROI, organizadas por clase, se integraron en una base de datos de imágenes, conservando sus dimensiones en píxeles para el análisis posterior.

II-B. Preparación de estructuras

Las redes neuronales convolucionales (RNC) se utilizan recientemente para la clasificación de imágenes, especialmente

aquellas con complejidades morfológicas y de color. Estos modelos han incrementado en complejidad y precisión mediante la modificación de sus arquitecturas, funciones de activación y estrategias de entrenamiento [19]. A continuación, se describen brevemente las arquitecturas de los modelos implementados:

- *GoogLeNet*: Desarrollada por investigadores de Google, ganó notoriedad al ganar el ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) en 2014. Se ha aplicado en detección de objetos, reconocimiento de escenas y teledetección [20], [21], [22], [23]. Esta RNC mejora la eficiencia de clasificación mediante el uso de unidades *Inception* y se implementa en software especializados como Matlab y Python, con alrededor de 144 capas de convolución. Para más detalles sobre esta arquitectura y su integración para la clasificación de imágenes, véase [20].
- *AlgarroboNet*: Esta estructura de RNC se propone para la clasificación de árboles de algarrobo *plus* y *no plus*, inspirada en la arquitectura de AlexNet (ver Figura 2). Consta de 18 capas, alternando cuatro capas de convolución entre las cuales se aplican capas de normalización por lotes, activación (ReLU) y MaxPool. Al finalizar las convoluciones, se utiliza una capa densa o completamente conectada que, mediante una función Softmax, permite la identificación de dos clases.

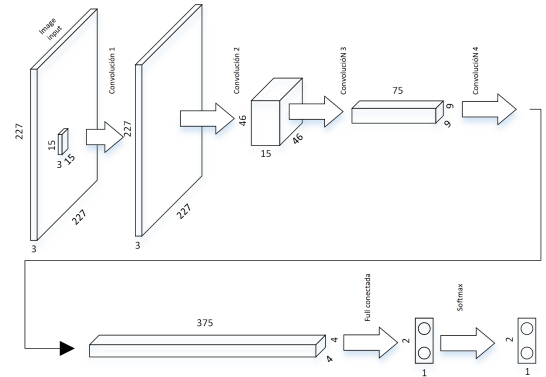


Figura 2: Esquema de AlgarroboNet

II-C. Entrenamiento de estructuras

Las imágenes de la base de datos se dividieron aleatoriamente siguiendo una estrategia *hold-out*, utilizando el 70 % para entrenamiento y el 30 % restante para validación. El entrenamiento de las RNC se realizó mediante la función *trainNetwork* de Deep Learning Toolbox™ en Matlab 2023a.

El proceso se repitió treinta veces, aleatorizando en cada oportunidad las imágenes a utilizar. Los parámetros de entrenamiento establecidos fueron:

- Optimizador = sgdm
- MiniBatchSize = 80
- MaxEpochs = 50
- InitialLearnRate = 1e-4
- Shuffle = every-epoch

- ValidationFrequency = 5
- ExecutionEnvironment = gpu

El entrenamiento se llevó a cabo en una computadora con procesador Intel(R) Xeon(R) Gold 5218 (2.30GHz), 96GB de RAM y sistema operativo Windows 10 Pro de 64 bits.

II-D. Evaluación de métricas

El rendimiento de cada prueba se evaluó mediante una matriz de confusión (CM) de dos clases, que resume la cantidad de verdaderos positivos (TP), verdaderos negativos (TN), falsos positivos (FP) y falsos negativos (FN), como se muestra en la Figura 3.

Clases predichas	cp ₂	TP	FP
	cp ₁	FN	TN
		cr ₁	cr ₂
		Clases reales	

Figura 3: Esquema de matriz de confusión Bi-clase

La CM organiza las clases verdaderas y predichas de los elementos de una muestra en una matriz cuadrada de tamaño $n \times n$, donde $n = 2$ para las clases *plus* y *non-plus*. A partir de la CM, se calcularon las siguientes métricas estándar [24], utilizadas en múltiples aplicaciones [25], [26]:

$$Accuracy \quad (ACC) = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Recall \quad (Rec) = \frac{TP}{TP + FN} \quad (2)$$

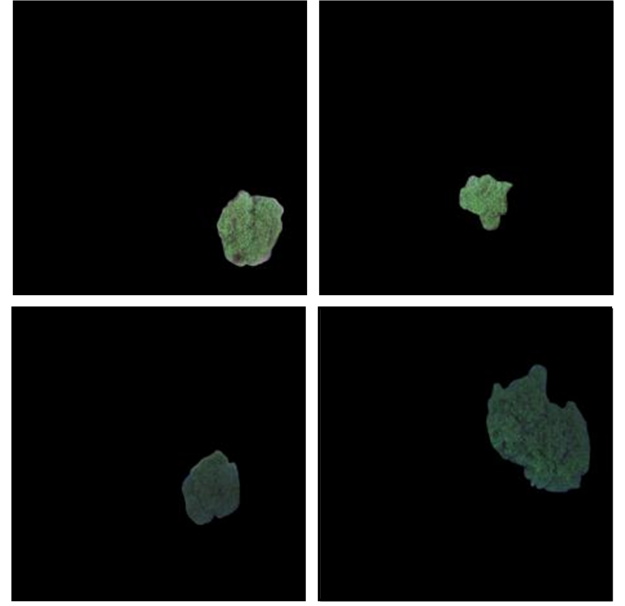
$$Precision \quad (Pre) = \frac{TP}{TP + FP} \quad (3)$$

$$F2_{measure} \quad (F2M) = 2 * \frac{Pre * Rec}{Pre + Rec} \quad (4)$$

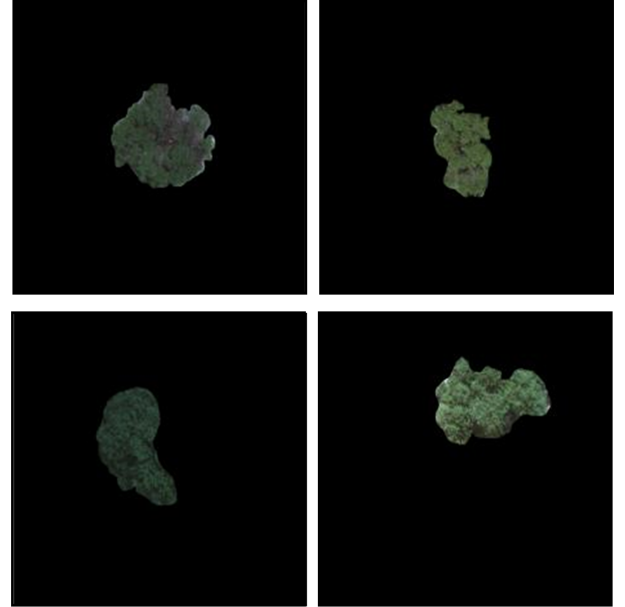
III. RESULTADOS Y DISCUSIÓN

III-A. Regiones de interés

Las ROI que forman parte de la base de datos de imágenes del proyecto, seleccionadas en las imágenes obtenidas mediante un VANT, se muestran en la Figura 4. Se puede observar una alta variabilidad en la forma y el color tanto dentro de una misma clase como entre diferentes clases. Esta variabilidad presenta un desafío significativo para la clasificación basada en la extracción manual de parámetros morfogeométricos y el entrenamiento de clasificadores tradicionales de machine learning, como las máquinas de soporte vectorial (MSV) y las redes neuronales de base radial (RNBR).



(a) No plus



(b) Plus

Figura 4: Ejemplos de ROI segmentadas en las imágenes obtenidas por VANT

El número de ROI fue de 485 para la clase “plus” y 490 para la clase “no plus”. Este número es inferior al recomendado para el entrenamiento de redes neuronales convolucionales (RNC), donde generalmente se sugieren al menos 1000 muestras por clase, como menciona Oblitas et al. [27]. Sin embargo, no es posible aplicar una estrategia de incremento de la base de datos mediante operadores morfológicos (escalado, rotación, translación, entre otros) debido a que la base del análisis se centra en las diferencias en sus patrones.

III-B. Preparación de estructuras

La preparación de las RNC involucró la generación de dos archivos *.mat que contienen la información de las capas (Layer). Los parámetros de entrenamiento de las RNC y algunos de los detalles más relevantes se presentan en la Tabla I.

TABLA I.: Principales parámetros de las redes en estudio

Neural Network	Total Layers	Convolutional Layers	Trainable parameters
AlgarroboNet	18	4	1 655 192
GoogleNet	144	59	6 998 552

El número de parámetros entrenables es 4.2 veces mayor en la red *GoogleNet* en comparación con la *AlgarroboNet*; sin embargo, como se discutirá a continuación, esto depende del tipo de estrategia de entrenamiento utilizada [28].

III-C. Resultados de entrenamiento

El proceso de entrenamiento, realizado bajo las condiciones establecidas y utilizando el sistema de cómputo descrito previamente, requirió un promedio de 365.24 segundos para la red *GoogleNet* y 722.61 segundos para la *AlgarroboNet*. Para comparar los tiempos de entrenamiento, se calcularon los tiempos relativos de los treinta entrenamientos, como se muestra en la Figura 5.

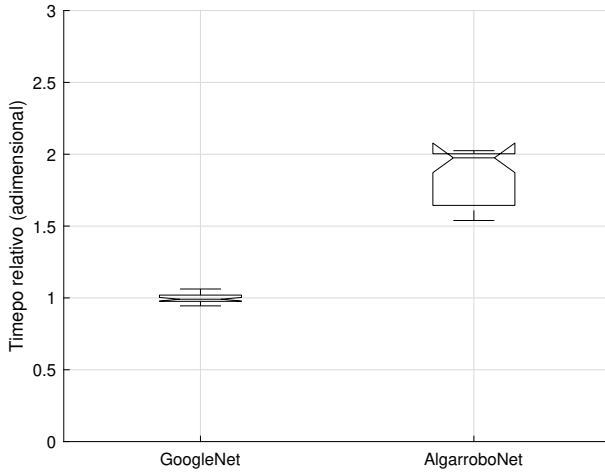


Figura 5: Tiempo relativo de entrenamiento de ambas RNC

El proceso de entrenamiento mostró que la red *AlgarroboNet* alcanzó estabilidad alrededor de las 350 iteraciones, como se observa en la Figura 6. Esto es superior a lo requerido por la red *GoogleNet*, que alcanzó estabilidad en las primeras iteraciones.

El entrenamiento de la red *GoogleNet* se realizó mediante una técnica de transferencia de aprendizaje o ajuste fino, mientras que la red *AlgarroboNet* se entrenó desde cero (*from scratch*), similar a los trabajos de Ávila et al. [26] y Castro et al. [28]. Por lo tanto, a pesar del mayor número de parámetros de la *GoogleNet*, el tiempo relativo de entrenamiento fue

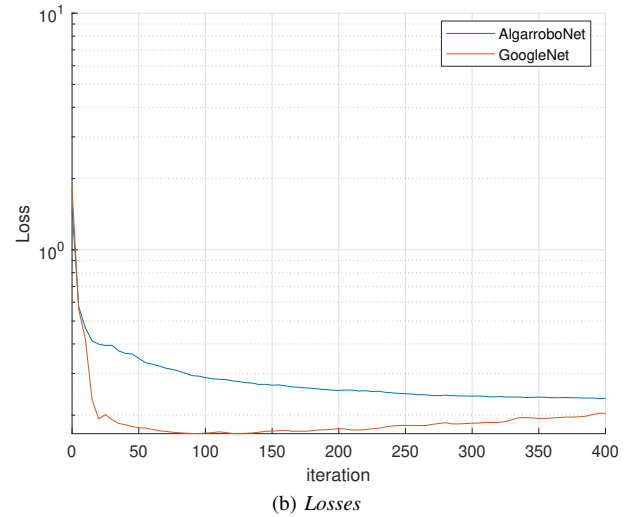
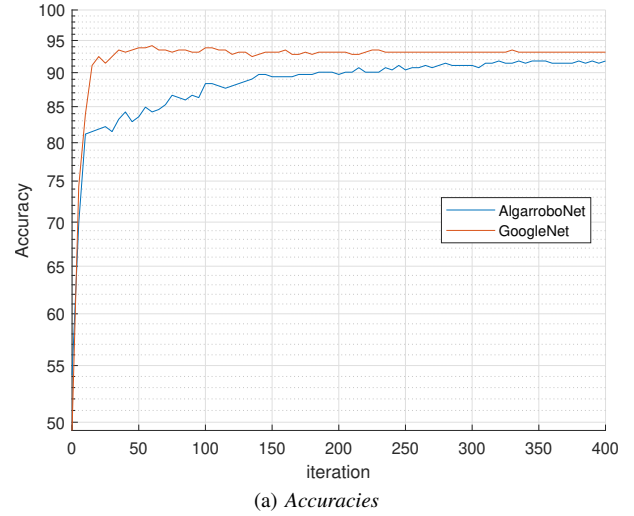


Figura 6: Gráficas promedio de entrenamientos de las RNC

similar debido a las diferentes estrategias de entrenamiento utilizadas.

III-D. Métricas de clasificación

Las métricas de los modelos de clasificación, mostradas en la Figura 7, indican que la precisión (*accuracy*) de la red *GoogleNet* (0.95) es ligeramente superior a la obtenida por la red *AlgarroboNet* (0.91). A pesar de la alta complejidad relativa de la *GoogleNet*, expresada en ocho veces el número total de capas y 4.2 veces el número de parámetros entrenables, la *AlgarroboNet* logró valores cercanos en términos de *accuracy* y $F1_{measure}$.

Este resultado es consistente con lo mencionado por Avila et al. [26], quienes compararon diferentes tipos de redes neuronales convolucionales y concluyeron que la complejidad de las redes debe evaluarse en función de la complejidad de la tarea a realizar.

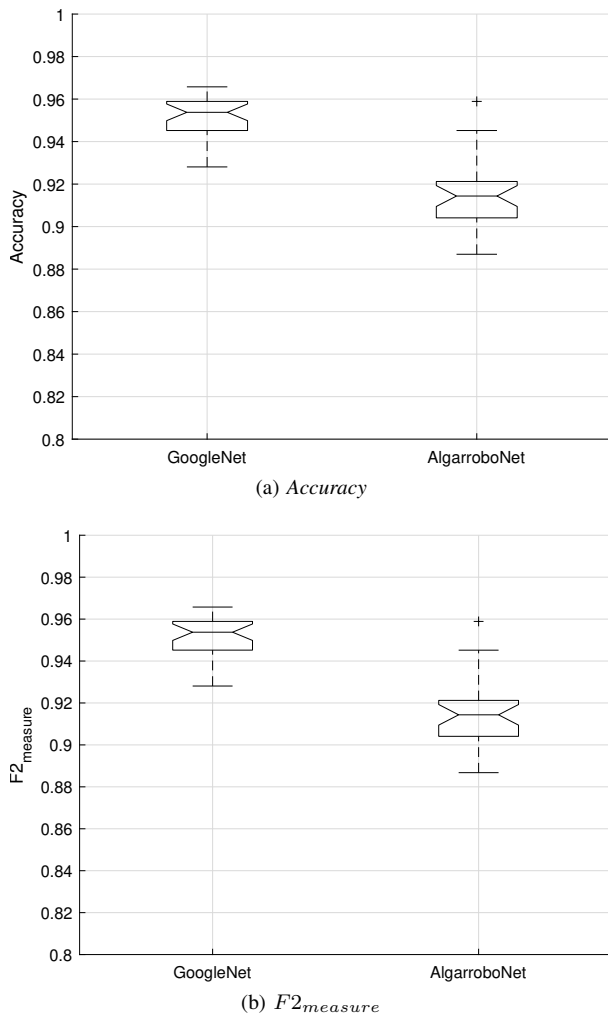


Figura 7: Métricas de clasificación

IV. CONCLUSIONES

El estudio evaluó dos RNC de diferente nivel de complejidad para la clasificación de árboles plus de algarrobo, extrayendo sus métricas de entrenamiento.

El entrenamiento de una red con hasta ocho veces el número de capas de convolución, mediante transferencia de aprendizaje, fue aproximadamente la mitad del tiempo requerido para la red de menor complejidad, entrenada desde cero. Comparativamente, la red de mayor complejidad solo mejoró en un 4% los resultados de la red más simple. Esto podría mejorarse aumentando el entrenamiento con una base de datos de mayor número de elementos.

Se debería desarrollar métricas de evaluación de la complejidad de las tareas de clasificación para establecer procedimientos claros que eviten el sobredimensionamiento de las RNC.

AGRADECIMIENTOS

Los autores agradecen el apoyo brindado a este trabajo por el proyecto “Redes convolucionales y combinaciones de bandas espectrales e índices vegetativos para identificación de

árboles plus de algarrobo (*Prosopis pallida*): aplicación de técnicas Deep learning e imágenes multispectrales” con contrato No. PE501078266-2022 y al proyecto “Creación del servicio de laboratorio de investigación en inocuidad de alimentos en la Universidad Nacional de Frontera”, con Código Único de Inversión N°243954.

REFERENCIAS

- [1] United Nations, *Biodiversity and Sustainable Development Goals*. New York, United States of America: Departamento de Asuntos Económicos y Sociales, 2023.
- [2] FAO, “The state of the world’s forests: Forests, biodiversity, and people [el estado de los bosques del mundo: Los bosques, la biodiversidad y las personas],” Centro de Información de Recursos Naturales (CIREN), Tech. Rep., 2020.
- [3] —, “The state of the world’s forests 2022: Forest pathways for green recovery and building inclusive, resilient, and sustainable economies [el estado de los bosques del mundo 2022. vías forestales hacia la recuperación verde y la creación de economías inclusivas, resilientes y sostenibles],” Centro de Información de Recursos Naturales (CIREN), Tech. Rep., 2022.
- [4] E. Sobrino, F. del Águila, and M. Torres, “Analysis of deforestation and vegetation loss at the national level and its regional impact [análisis de la deforestación y pérdida de vegetación a nivel nacional y el impacto a nivel regional],” Centro Nacional de Planeamiento Estratégico - CEPLAN, Tech. Rep., 2023.
- [5] A. Pommerening, K. Maleki, and J. Haufe, “Tamm review: Individual-based forest management or seeing the trees for the forest,” *Forest Ecology and Management*, vol. 501, p. 119677, 2021.
- [6] M. Dieng, C. Mbow, D. L. Skole, and B. Ba, “Sustainable land management policy to address land degradation: linking old forest management practices in senegal with new redd+ requirements,” *Frontiers in Environmental Science*, 2023.
- [7] M. Weiss, F. Jacob, and G. Duveiller, “Remote sensing for agricultural applications: A meta-review,” *Remote sensing of environment*, vol. 236, p. 111402, 2020.
- [8] E. Ponce-Corona, M. G. Sánchez, D. Fajardo-Delgado, B. Acevedo-Juárez, M. De-la Torre, H. Avila-George, and W. Castro, “A systematic literature review focused on the use of unmanned aerial vehicles during vegetation detection processes [una revisión sistemática de la literatura enfocada al uso de vehículos aéreos no tripulados durante el proceso de detección de vegetación],” *RISTI - Revista Ibérica de Sistemas e Tecnologías de Informação*, no. 36, pp. 82–101, Mar. 2020.
- [9] M. Gouvêia, M. Melo, R. Augusti, V. Silva, Y. Lobo, Y. Figueiredo, J. Melo, K. Canteral, J. Viana, and K. Ferreira, “Procesos sucesionales de la vegetación bajo plantaciones forestales y bosque andino en la hoya de loja, ecuador,” *Conservação e biodiversidade amazônica: potencialidade e incertezas*, vol. 2, no. 1, pp. 107–127, 2023.
- [10] R. P. Sishodia, R. L. Ray, and S. K. Singh, “Applications of remote sensing in precision agriculture: A review,” *Remote sensing*, vol. 12, no. 19, p. 3136, 2020.
- [11] B. Chehreh, A. Moutinho, and C. Viegas, “Latest trends on tree classification and segmentation using uav data—a review of agroforestry applications,” *Remote Sensing*, vol. 15, no. 9, p. 2263, 2023.
- [12] D. C. Tsouros, S. Bibi, and P. G. Sarigiannidis, “A review on uav-based applications for precision agriculture,” *Information*, vol. 10, no. 11, p. 349, 2019.
- [13] L. Benos, A. C. Tagarakis, G. Dolias, R. Berruto, D. Kateris, and D. Bochtis, “Machine learning in agriculture: A comprehensive updated review,” *Sensors*, vol. 21, no. 11, p. 3758, 2021.
- [14] B. Debus, H. Parastar, P. Harrington, and D. Kirsanov, “Deep learning in analytical chemistry,” *TrAC Trends in Analytical Chemistry*, vol. 145, p. 116459, 2021.
- [15] F. Zhu, Z. Chen, H. Li, Q. Shi, and X. Liu, “Cedonet: Individual tree segmentation in dense orchard via context enhancement and density prior,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 7040–7051, 2024.
- [16] R. Ipinza, “Methods for selecting plus trees [métodos de selección de árboles plus],” Biblioteca digital Instituto Forestal, Tech. Rep., 1998.

- [17] H. Avila-George, T. Valdez-Morones, H. Pérez-Espinosa, B. Acevedo-Juárez, and W. Castro, "Using artificial neural networks for detecting damage on tobacco leaves caused by blue mold," *Strategies*, vol. 9, no. 8, 2018.
- [18] W. Castro, J. Oblitas, J. Maicelo, and H. Avila-George, "Evaluation of expert systems techniques for classifying different stages of coffee rust infection in hyperspectral images," *International Journal of Computational Intelligence Systems*, vol. 11, no. 1, pp. 86–100, 2018.
- [19] L. Chen, S. Li, Q. Bai, J. Yang, S. Jiang, and Y. Miao, "Review of image classification algorithms based on convolutional neural networks," *Remote Sensing*, vol. 13, no. 22, p. 4712, Nov. 2021.
- [20] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.
- [21] P. Tang, H. Wang, and S. Kwong, "G-ms2f: Googlenet based multi-stage feature fusion of deep cnn for scene recognition," *Neurocomputing*, vol. 225, pp. 188–197, Feb. 2017.
- [22] Z. Zhang, P. Flores, C. Igathinathane, D. L. Naik, R. Kiran, and J. K. Ransom, "Wheat lodging detection from uas imagery using machine learning algorithms," *Remote Sensing*, vol. 12, no. 11, p. 1838, Jun. 2020.
- [23] W. Gomez-Flores, J. J. Garza-Saldana, and S. E. Varela-Fuentes, "A huanglongbing detection method for orange trees based on deep neural networks and transfer learning," *IEEE Access*, vol. 10, pp. 116 686–116 696, 2022.
- [24] T. Fawcett, "An introduction to roc analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, Jun. 2006.
- [25] S. Ruuska, W. Hämäläinen, S. Kajava, M. Mughal, P. Matilainen, and J. Mononen, "Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle," *Behavioural Processes*, vol. 148, pp. 56–62, Mar. 2018.
- [26] H. Avila-George, M. De-la Torre, J. Sánchez-Garcés, J. J. C. Quispe, J. M. Prieto, and W. Castro, "Discrimination of foreign bodies in quinoa (*chenopodium quinoa* willd.) grains using convolutional neural networks with a transfer learning approach," *PeerJ*, vol. 11, p. e14808, 2023.
- [27] J. Oblitas, J. Mejia, M. De-la Torre, H. Avila-George, L. Seguí Gil, L. Mayor López, A. Ibarz, and W. Castro, "Classification of the microstructural elements of the vegetal tissue of the pumpkin (*cucurbita pepo* l.) using convolutional neural networks," *Applied Sciences*, vol. 11, no. 4, p. 1581, 2021.
- [28] W. Castro, H. Yoshida, L. S. Gil, L. M. López, J. Oblitas, M. De-la Torre, and H. Avila-George, "Microstructural analysis in foods of vegetal origin: an approach with convolutional neural networks," in *2019 8th international conference on software process improvement (CIMPS)*. IEEE, 2019, pp. 1–5.