

Determinación del mejor algoritmo de reconocimiento de imágenes de camélidos sudamericanos mediante machine learning, Apurímac, 2023

Determination of the best image recognition algorithm for south american camelids using machine learning, Apurímac, 2023

Filio Carrasco-Sauñe ^A, Manuel J. Ibarra-Cabrera ^BORCID:0009-0006-5867-275X ^A, ORCID:0000-0001-6711-4916 ^B

(Recepción: 15/04/2023 y aceptación 08/10/2023)

Resumen— Las alpacas y las vicuñas pertenecen a la familia de los camélidos sudamericanos, la alpaca es un animal doméstico, mientras que la vicuña es un animal silvestre que generalmente vive en alturas superiores a los 3000 m.s.n.m. El problema es que en el caso de las vicuñas es difícil realizar el reconocimiento y conteo y porque por su naturaleza de ser animal silvestre huyen de las personas y de cualquier otro objeto extraño que no sea de su entorno; el otro problema es que no se tiene una base de datos de imágenes o fotos de vicuñas para hacer un estudio de reconocimiento de imágenes. Este trabajo consistió en recolectar 146 imágenes de vicuña, de las cuales 95, fueron consideradas para el entrenamiento 41 para la validación y 10 para las pruebas; las imágenes fueron recolectadas en un centro poblado del distrito de Cotaruse, provincia de Aymaraes en la región Apurímac de Perú. Para el caso de las alpacas, las imágenes se obtuvieron del repositorio de Kaggle y se trabajó con 142 imágenes de las cuales 95 fueron consideradas para el entrenamiento, 41 para la validación y 10 para las pruebas. Los resultados muestran que el algoritmo o modelo Mask R-CNN obtiene valor para accuracy de 1,0 para las vicuñas y de 0,95 para las alpacas; estos valores son los más eficientes en relación con los encontrados por Yolo V8 y SSMD.

Palabras clave: camélidos, reconocimiento, red neuronal, vicuña

Abstract— Alpacas and vicuñas belong to the South American camelid family; the alpaca is a domestic animal, while the vicuña is a wild animal that generally lives at altitudes above 3,000 meters above sea level. The problem is that in the case of vicuñas it is difficult to recognize and count them because they are wild animals and because they flee from people and any other foreign object that is not part of their environment; the other problem is that there is no database of images or photos of vicuñas to carry out an image recognition study. This work consisted of collecting 146 vicuña images, of which 95 were considered for training, 41 for validation and 10 for testing; the images were collected in a village center in the district of Cotaruse, province of Aymaraes in the Apurímac region of Peru. In the case of alpacas, the images were obtained from the Kaggle repository and 142 images were used, of which 95 were considered for training, 41 for validation and 10 for testing. The results show that the Mask R-CNN algorithm or model obtains a value for accuracy of 1.0 for vicuñas and 0.95 for alpacas; these values are the most efficient in relation to those found by Yolo V8 and SSMD.

Keywords: camelids, neural network, recognition, vicuña

A. Filio Carrasco-Sauñe, Escuela Profesional de Ingeniería Informática y Sistemas de la Universidad Nacional Micaela Bastidas de Apurímac-Perú, 112134@unamba.edu.pe.

B. Manuel J. Ibarra-Cabrera, Departamento Académico de Informática y Sistemas de la Universidad Nacional Micaela Bastidas de Apurímac-Perú, mibarra@unamba.edu.pe.

1 INTRODUCCIÓN

Actualmente, existen cuatro especies de camélidos sudamericanos bastante populares: llamas y alpacas que son conocidos como animales domésticos; y vicuñas y guanacos que son conocidos como animales silvestres.

La vicuña es uno de los animales que vive en los andes del Perú y resulta complicado determinar con precisión la cantidad exacta de estos seres vivos. Esto se debe a que llevar a cabo el monitoreo de la población de animales silvestres resulta difícil y todavía se basa en el conteo manual o en el

uso de transectos, que es un método de muestreo para calcular la densidad poblacional.

Las cuatro especies de los camélidos sudamericanos son: llamas y alpacas que son conocidos como animales domésticos; y vicuñas y guanacos que son conocidos como animales silvestres [1]. La vicuña es una especie animal silvestre que está en peligro de extinción, debido a que su lana es de alta calidad y tiene un costo elevado en el mercado [2], por tanto, parece ser atractivo en término de costos para las personas que los cazan furtivamente [3]. Mayormente viven sobre los 3000 m.s.n.m. en la zona sierra o los andes de los países de Argentina, Bolivia, Chile, Ecuador y Perú [4].

El problema es que realizar el conteo de animales silvestres es un proceso complejo, específicamente el conteo de vicuñas [5] se realiza una vez al año o cada dos años en una actividad que se llama “vicuña chaco”, y se usa el método conteo por aproximación, o transecto [6]. Para realizar el conteo, previamente se requiere el reconocimiento, por esta razón, esta investigación se enfoca en el reconocimiento de imagen de vicuña.

Un segundo problema asociado al conteo de las vicuñas, es que no existen imágenes de vicuñas disponibles en las bases de datos de libre uso y acceso para realizar reconocimiento y conteo utilizando algoritmos de machine learning. Una alternativa para obtener las imágenes de animales silvestres, es uso de un dron o cámaras fotográficas escondidas, y así tener una opción de realizar el reconocimiento, conteo y monitoreo de animales [7].

Por otro lado, se tiene disponible una base de datos de imágenes para la clasificación de alpacas en *Kaggle* [8] (repositorio de imágenes para realizar entrenamientos y pruebas con algoritmos de machine learning); sin embargo, no se ha podido encontrar base de datos para llama, guanaco ni vicuña, que también son especies que pertenecen a los camélidos sudamericanos, esto dificulta realizar entrenamiento y pruebas con algoritmos de machine learning.

Los algoritmos de machine learning cada vez son más importantes y se pueden aplicar a diversas áreas, consiste en entrenar un conjunto de imágenes en base ciertas reglas o datos, después se tiene un modelo entrenado y finalmente se realizan las pruebas para determinar la precisión [9]. Existen algoritmos para el reconocimiento de objetos, y están basados en redes neuronales, redes bayesianas y basados en características [10], [11].

Existen algunos autores que han estudiado el reconocimiento de animales, así por ejemplo Xiao et al. [12] realizaron una investigación cuyo objetivo fue proponer un método para la identificación o reconocimiento de vacas basado en visión por computadora. La metodología que siguieron fue obtener las imágenes en un ambiente libre en un granero, debido a los cambios en la posición y el ángulo de la cámara, los patrones de las vacas grabados a menudo se deformaron, lo que dificultó la identificación individual de las vacas. Primero, se obtiene una imagen de vista superior de una vaca y se utiliza una Mask R-CNN mejorada para segmentar esta imagen y extraer las características de la forma del lomo de la vaca. Luego, se utilizó un enfoque de Fisher para seleccionar el mejor subconjunto de características y se aplicó un clasificador de máquina de vectores de soporte (SVM) para identificar vacas individuales. Para verificar la efectividad del algoritmo de detección de objetivos, el método propuesto se compara con el modelo Mask R-CNN tradicional, y se comparan la precisión, la recuperación, la puntuación F1, el tiempo de ejecución promedio por imagen y la precisión promedio del modelo Mask R-CNN mejorado. 98,21%, 96,48%, 97,34%, 1,02s y 97,39%, respectivamente. Para la

identificación individual de las vacas se utiliza un clasificador SVM entrenado en función de las características de forma obtenidas. El método propuesto logra una precisión de identificación de vacas del 98,67% basándose en un conjunto de datos que contiene imágenes de vista superior de 48 vacas. Los resultados demuestran la efectividad del método de identificación de vacas propuesto y su importante potencial para su uso en el manejo de precisión de las vacas lecheras.

Por otro lado, Li et al. [13] realizaron una investigación que usa el aprendizaje profundo para reconocer las características de vacas individuales. Alexnet se utiliza como red esqueleto para una red neuronal convolucional liviana que puede reconocer vacas individuales en imágenes con fondos complejos. Se agregó un módulo de inicio mejorado y un mecanismo de atención para extraer características en múltiples escalas para mejorar la detección de puntos característicos. En los experimentos, se recogieron imágenes de vista lateral de 13 vacas. El método propuesto logró una precisión del 97, 95% en la identificación de vacas con un único tiempo de entrenamiento de solo 6s, que es una sexta parte del Alexnet original. Para verificar la validez del modelo, el conjunto de datos y los parámetros experimentales se mantuvieron constantes y se compararon con los resultados de Vgg16, Resnet101, Mobilenet V2 y GoogLeNet. El modelo propuesto garantiza una alta precisión y tiene el tamaño de parámetro más pequeño de 6,51 MB, que es 1,3 veces menor que el de la red Mobilenet V2, famosa por su peso ligero.

El objetivo de este proyecto de investigación fue determinar el mejor algoritmo de Machine Learning para el reconocimiento de dos especies de los camélidos sudamericanos; específicamente en el reconocimiento de imágenes de vicuñas, para lo cual se recolectó imágenes de Apurímac mediante cámara fotográfica; y por otro lado, en el reconocimiento de imágenes de alpacas de acuerdo al dataset de la base de datos en Kaggle. El entrenamiento se realizó con tres modelos o algoritmos de redes neuronales convolucionales: Yolo V8 [14], Mask R-CNN Resnet101 [15] y SSMD (Single Shot Multibox Detector (SSMD) mobileNet 2.0 [16].

2 MÉTODO

2.1 Ámbito de estudio

Este estudio se realizó en la Universidad Nacional Micaela Bastidas de Apurímac (UNAMBA) en la ciudad de Abancay; y los datos fueron recolectados en el distrito de Cotaruse y provincia de Aymaraes en el año 2023.

2.2 Tipo y nivel de investigación

Este trabajo es una investigación de tipo aplicada y de nivel descriptivo.

2.3 Procedimiento. fue el siguiente:

- a) Se recolectó las imágenes de vicuñas mediante cámara fotográfica de un celular *Iphone 11 pro max*, con una resolución máxima de 2268x4032 pixeles, y la calidad de

video de 30 fotogramas por segundo.

- Se eliminaron las imágenes que estuvieran desenfocadas, con mala calidad o que tengan muchas imágenes superpuestas.
- Se separó aproximadamente 70,0% imágenes para entrenamiento, 20,0% para validación y 10,0% de las imágenes para pruebas.
- Luego, sólo a las imágenes separadas para el entrenamiento se procedió a etiquetarlas utilizando el software de etiquetado CVAT [17], que es una plataforma especializada para realizar anotaciones sobre imágenes.
- Se descargó de *Kaggle* las imágenes de alpaca y se procedió a entrenar con las diferentes técnicas de machine learning.
- Luego, se entrenó las imágenes de vicuña y alpaca con las diferentes técnicas de machine learning, y se utilizó un *cuaderno de Google Collab* con la cuenta de correo electrónico de *filiocs27@gmail.com*. El entrenamiento demoraba alrededor de 40 minutos con una cuenta estándar de Google, lo cual no permitía completar el proceso de entrenamiento; por lo que se tuvo que adquirir y utilizar una cuenta *premium de Google* para que el entrenamiento no se corte, esto se consigue pagando 9,90 dólares adicionales. El entrenamiento se realizó utilizando herramientas de programación y librerías con sus distintas versiones, por ejemplo: `cuda-toolkit==11.0.3`, `cuda-nn==8.0.5.39`, `numpy==1.19.2`, `scikit-learn==1.2.0`, `scipy==1.10.1`, `seaborn==0.10.0`, `pandas==2.0.3`, `pillow==10.0.0`, `matplotlib==3.6.0`, `tensorflow-gpu==2.4.0`, `opencv-python==4.5.4.60`, `ipython==8.10.0`.
- Luego se realizaron las pruebas con cada una de las técnicas con las imágenes separadas para el testing (imágenes sin etiquetar).

2.4 Población

Las imágenes o fotos de vicuñas se obtuvieron en el distrito de Cotaruse provincia de Aymaraes en la región Apurímac, cuyas coordenadas geográficas son 14°31'43.4"S y 73°27'15.6"W. Se obtuvieron un total de 146 imágenes de buena calidad. La Fig. 1 muestra las imágenes tomadas en los andes de Apurímac a más de 3000 m.s.n.m.



Fig. 1. Imágenes tomadas en el campo

Para el proceso de entrenamiento se obtuvo una muestra de 96 imágenes para entrenamiento, 41 imágenes para la validación y 10 para el testing. Este muestreo fue de forma aleatoria.

Para el caso de las alpacas la población fue extraída del repositorio de imágenes *Kaggle*, con un total de 142 imágenes de alpacas; de las cuales 102 se utilizaron para el entrenamiento, 30 para validación, y 10 para el testing.

2.5 Materiales e Instrumentos

Dentro de los materiales utilizados están: a) Cámara fotográfica b) Google Collab; y c) Recurso repositorio digital *Kaggle*.

Dentro de los instrumentos utilizados están: a) Registro de imágenes de vicuñas; b) Cuaderno notebook cuyo archivo tiene extensión *.ipynb; y c) Registro de imágenes de Alpacas

3 RESULTADOS Y DISCUSIÓN

3.1 Pruebas y resultados para las vicuñas

La tabla 1 muestra la Matriz de Confusión para las vicuñas, en la cual se puede observar los resultados de los tres modelos entrenados, Yolo V8, Mask R-CNN Resnet101 y SSMD mobileNet 2.0, "TP" representa a los verdaderos positivos, "FP" los falsos positivos "FN" falsos negativos y "TN" verdaderos negativos.

TABLA 1

Matriz de confusión para vicuña

Algoritmo o modelo de Machine Learning	TP	FP	FN	TN
Yolo V8	11	0	2	0
Mask R-CNN Resnet101	11	0	0	0
SSMD mobileNet 2.0	10	1	2	0

La tabla 2 muestra los resultados estadísticos de las pruebas realizadas con las imágenes de vicuñas en los modelos o algoritmos de Yolo V8, Mask R-CNN Resnet101 y SSMD mobileNet 2.0, asimismo se tienen tres variables de medición "Accuracy", "Precision" y "Recall". Los resultados muestran que Mask R-CNN Resnet101 tiene el mejor resultado con un valor de 100%, luego esta Yolo V8 y finalmente SSMD. Estos resultados fueron realizados considerando 200 épocas de entrenamiento.

TABLA 2

Estadísticas para vicuña

Algoritmo o modelo de Machine Learning	Accuracy	Precision	Recall
Yolo V8	1,00	1,00	0,85
Mask R-CNN Resnet101	1,00	1,00	1,00
SSMD mobileNet 2.0	0,92	0,90	0,83

La tabla 3 muestra las predicciones realizadas a un conjunto de 10 fotos de imágenes de vicuñas, donde "N" significa un pseudónimo asignado a la figura, "cantidad" significa la cantidad de vicuñas que hay en la foto, y seguidamente están los modelos: Yolo V8, Mask R-CNN y SSMD. Así por ejemplo, la figura con pseudónimo "V10", contiene 2 imágenes de vicuñas, y Yolo V8 obtiene un valor 0,803 y 0,94; Mask R-CNN 0,997 y 0,997; y SSMD 0,65 y 0,94. Por lo que se puede observar que Mask R-CNN obtiene mejores resultados.

TABLA 3
Predicciones para Vicuña

N	Cantidad	Yolo V8	Mask R-CNN Resnet101	SSMD mobileNet 2.0
V01	1	0,92	0,997	0,89
V02	1	0,92	0,996	0,90
V03	1	0,92	0,999	0,89
V04	1	0,96	0,998	0,88
V05	1	0,93	0,998	0,86
V06	1	0,90	0,998	0,87
V07	1	0,84	0,999	0,44
V08	1	0,91	1,000	9,83
V09	1	0,89	0,999	0,79
V10	2	0,83-0,94	0,997-0,997	0,65-0,94

La Fig. 2 muestra el resultado de las predicciones en las vicuñas, en la que se observa que Mask R-CNN obtiene los mejores valores con 0,997 y 0,997 para cada vicuña.

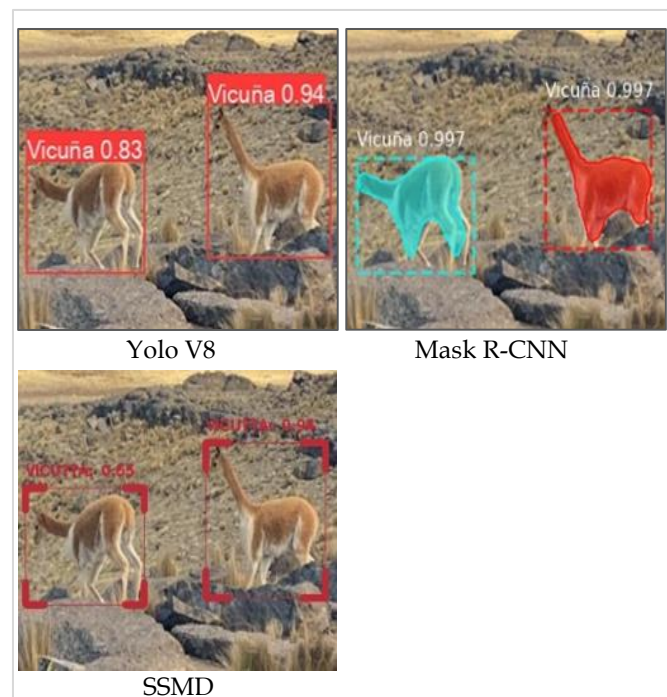


Fig. 2. Predicciones para imágenes de vicuña.

Para el caso de Mask R-CNN, la Fig. 3 muestra las métricas de precisión y recall versus las épocas, y lo cual nos indica que a partir de la época 60 aproximadamente ya se muestra un valor estabilizado en el proceso de entrenamiento; de forma similar, la Fig. 4 muestra la pérdida versus las épocas, y lo cual nos indica a partir de la época 80 ya se muestra un valor estabilizado en el proceso de entrenamiento.

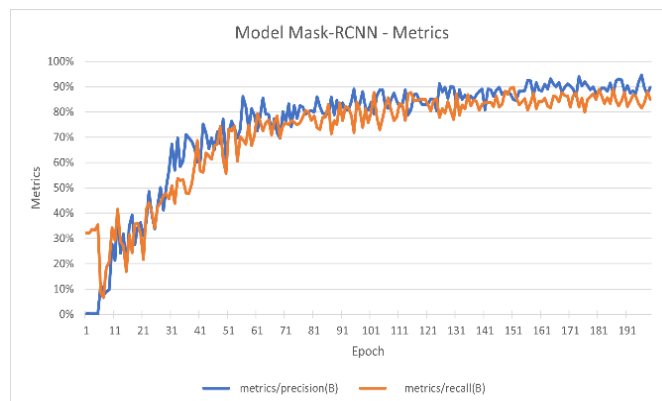


Fig. 3. Métricas vs época en Mask R-CNN

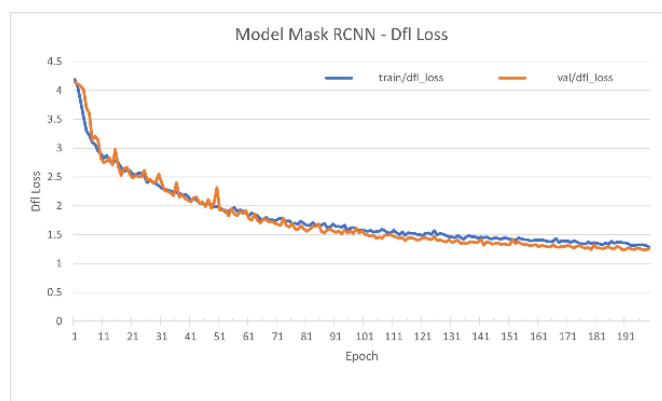


Fig. 4. Pérdida vs época en Mask R-CNN

Para el caso de Yolo V8, la Fig. 5 muestra las métricas de precisión y recall versus las épocas, y lo cual nos indica que a partir de la época 75 aproximadamente ya se muestra un valor estabilizado en el proceso de entrenamiento; de forma similar, la Fig. 6 muestra la pérdida versus las épocas, y lo cual nos indica que en la época 125 aproximadamente ya se muestra un valor estabilizado en el proceso de entrenamiento.

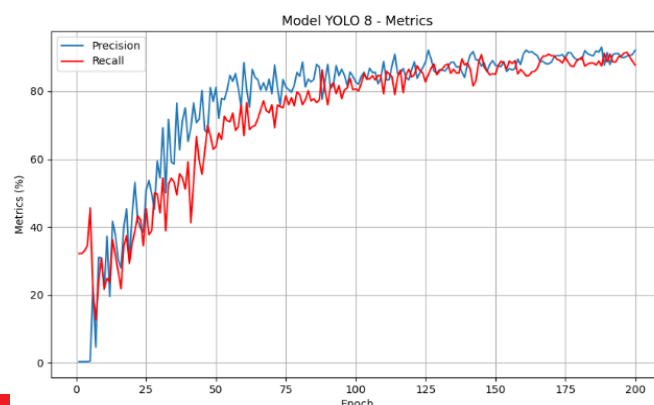


Fig. 5. Métricas vs época en Yolo V8

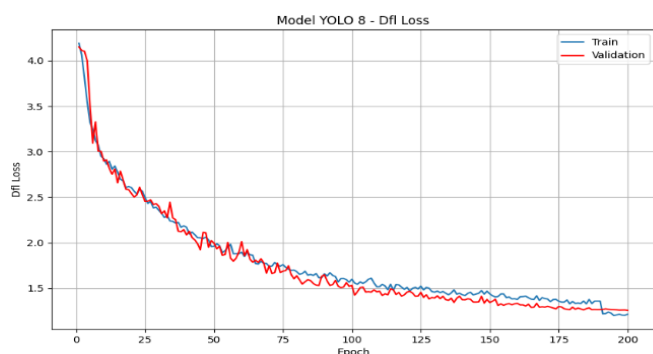


Fig. 6. Pérdida vs época en Yolo V8

3.2 Pruebas y resultados para las alpacas

La tabla 4 muestra la Matriz de Confusión para las alpacas, en la cual se puede observar los resultados de los tres modelos entrenados, Yolo V8, Mask R-CNN Resnet101 y SSMD mobileNet 2.0, "TP" representa a los verdaderos positivos, "FP" los falsos positivos "FN" falsos negativos y "TN" verdaderos negativos.

TABLA 4
Matriz de confusión para alpaca

Algoritmo o modelo de Machine Learning	TP	FP	FN	TN
Yolo V8	34	0	4	0
Mask R-CNN Resnet101	21	0	1	0
SSMD mobileNet 2.0	26	0	10	0

La tabla 5 muestra la Matriz de Confusión para las Alpacas, en la cual se puede observar los resultados de los tres modelos entrenados, Yolo V8, Mask R-CNN Resnet101 y SSMD mobileNet 2.0, "TP" representa a los verdaderos positivos, "FP" los falsos positivos "FN" falsos negativos y "TN" verdaderos negativos.

TABLA 5
Estadísticas para alpaca

Algoritmo o modelo de Machine Learning	Accuracy	Precision	Recall
Yolo V8	0,89	1,00	0,89
Mask R-CNN Resnet101	0,95	1,00	0,95
SSMD mobileNet 2.0	0,72	1,00	0,72

La tabla 6 muestra las predicciones realizadas a un conjunto de 10 fotos de imágenes de alpacas, donde "N" significa un pseudónimo asignado a la figura, "cantidad" significa la cantidad de alpacas que hay en la foto, y seguidamente están los modelos: Yolo V8 Mask R-CNN y SSMD. Así, por ejemplo, la figura "A10", contiene 3 imágenes de alpacas, y Yolo V8 obtiene un valor 0,57, 0,71 y 0,27; Mask R-CNN 0,998, 0,996 y

0,953; y SSMD 0,68, 0,85 y 0,56. Por lo que se puede observar que Mask R-CNN tiene mejores resultados

La Fig. 7 muestra el resultado de las predicciones en las alpacas, en la que se observa que Mask R-CNN obtiene el mejor valor de 0,999; esto implica que existe un 99.9% de certeza en que la imagen sea alpaca; en segundo lugar, se encuentra la predicción con Yolo V8 que tiene un valor de 0,84 y en tercer lugar se encuentra la predicción realizada con SSMD con un valor de 0,70.

TABLA 6
Predicciones para Alpaca

N	Cant idad	Yolo V8	Mask R-CNN Resnet101	SSMD mobileNet 2.0
A01	1	0,84	0,999	0,70
A02	1	0,84	0,999	0,53
A03	2	0,59-0,83	0,996-0,997	0,70-0,87
A04	2	0,97-0,98	0,999-0,988	0,81-0,88
A05	2	0,91-0,98	1,000-0,998	0,70-
A06	2	0,93-0,94	0,999-0,998	0,64 0,54
A07	1	0,91	1,000	0,68
A08	2	0,84-0,87	1,000-0,999	0,86-
A09	3	0,82-0,82	1,000-0,999-	0,81-0,75-0,75
A10	3	0,57-0,71-0,27	0,998-0,996-0,953	0,68-0,85-0,56

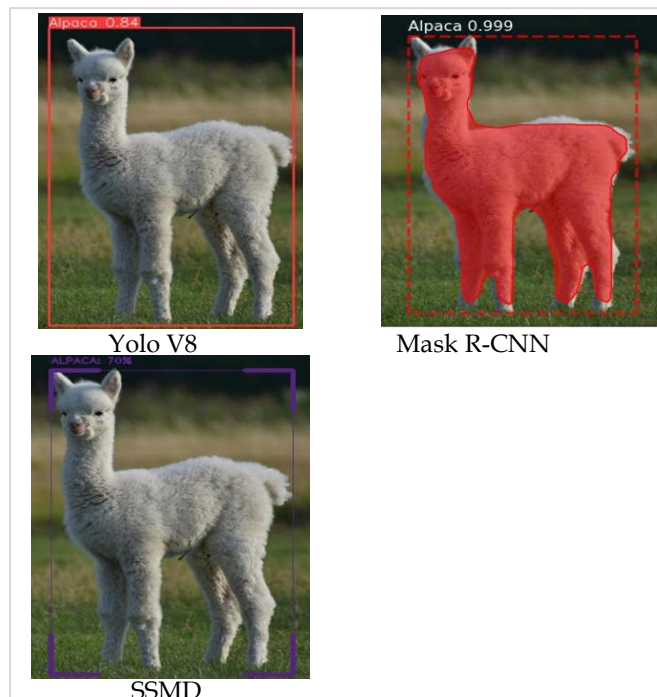


Fig. 7. Predicciones para imágenes de alpaca

3.3 Discusión

Según Xiao et al. [12], el reconocimiento de imágenes de 48 vacas lecheras mediante el modelo Mask R-CNN tradicional obtuvieron resultados satisfactorios con una precisión de 0,98, y la recuperación de 0,96 en una escala de 0 a 1. Para la identificación individual de las vacas se utiliza un clasificador SVM entrenado en función de las características

de forma obtenidas. Estos resultados están alineados con los resultados obtenidos por Xiao et al. [12], puesto que para el caso de las vicuñas se obtuvieron el modelo de Mask R-CNN valores de precisión 1,00 y recuperación de 0,95, y para el caso de las alpacas precisión de 1,00 y recuperación de 0,95. Es preciso mencionar, que no existe pruebas realizadas con camélidos sudamericanos, pero si existen investigación con vacas. Por otro lado, lo investigado por Li et al. [13] en el cual usan el aprendizaje profundo para reconocer las características de vacas individuales. En los experimentos muestran que, el método propuesto tiene una precisión del 97,95% en la identificación de vacas lecheras y con un único tiempo de entrenamiento de solo 6 s, que es una sexta parte del Alexnet original y que es 1,3 veces menor que el de la red Mobilenet V2, famosa por su peso ligero. La investigación de Li et al. [13] se enfoca prioritariamente en el tiempo de respuesta en la identificación, mientras que nuestra investigación es orientada para camélidos sudamericanos enfocado en el reconocimiento de imágenes y no en el tiempo de respuesta.

4 CONCLUSIONES

Este trabajo de investigación muestra que, para el algoritmo más eficiente (en base a la variable precisión) para el reconocimiento de camélidos sudamericanos, es Mask R-CNN Resnet101, seguido de Yolo V8 y finalmente SSMD. Los algoritmos fueron probados con 146 imágenes de vicuñas tomadas en Apurímac-Perú y 142 imágenes de alpacas extraídas de Kaggle.

AGRADECIMIENTOS

Queremos agradecer a CONCYTEC-FONDECYT por el apoyo en el desarrollo de esta investigación, mediante el fondo concursable que tiene contrato número 376-2019.

REFERENCIAS

- [1] H. D. Yacobaccio, "The domestication of South American camelids: a review," *Anim. Front.*, vol. 11, no. 3, pp. 43–51, 2021, <https://doi.org/10.1093/af/vfaa065>
- [2] J. E. Quispe Coaquira, T. Herrera Mamani, E. Apaza Zúñiga, L. Clavetea Quisca, and Z. Maquera Marón, "Características tecnológicas de la fibra de vicuñas en semicautiverio de la Multicomunal Picotani-Región Puno," *Rev. Investig. Vet. del Perú*, vol. 29, no. 2, pp. 522–532, 2018, <https://doi.org/10.15381/rivep.v29i2.14475>
- [3] G. Rodríguez García, "Animales...? En peligro de extinción o en peligro de que los extingamos?," *Ius Prax.*, vol. 14, no. 1, pp. 301–315, 2008, doi: <http://dx.doi.org/10.4067/S0718-00122008000100011>.
- [4] J. Laker, J. Baldo, Y. Arzamendia, and H. D. Yacobaccio, "La vicuña en los Andes," *Investig. Conserv. y manejo vicuñas*, pp. 37–50, 2006, [Online]. Available: <https://www.vicam.org.ar/publi/VicuniaAndes.pdf>
- [5] R. Hoffmann, K. Otte, and C. Ponce, *El manejo de la vicuña silvestre Tomo I*, vol. I. Alemania: Sociedad Alemana de Cooperación Técnica- GTZ, 1983.
- [6] G. Gallardo, A. Nuñez, and L. F. Pacheco, "Transectos lineales como opción para estimar abundancia de vicuñas (Vicugna vicugna): Estudio de caso en el Parque Nacional Sajama,

- Boivia," *Ecol. en Bolív.*, vol. 45, pp. 64–72, 2010.
- [7] K. S. Christie, S. L. Gilbert, C. L. Brown, M. Hatfield, and L. Hanson, "Unmanned aircraft systems in wildlife research: Current and future applications of a transformative technology," *Front. Ecol. Environ.*, vol. 14, no. 5, pp. 241–251, 2016, doi: [10.1002/fee.1281](https://doi.org/10.1002/fee.1281), <https://doi.org/10.1002/fee.1281>
- [8] Kaggle, "Image Dataset," 2023. <https://www.kaggle.com/datasets/sid4sal/alpaca-dataset-small> (accessed Jan. 02, 2023).
- [9] J. Hu, X. Wu, and M. Dai, "Estimating the population size of migrating Tibetan antelopes *Pantholops hodgsonii* with unmanned aerial vehicles," *Oryx*, pp. 1–9, 2018, <https://doi.org/10.1017/S0030605317001673>
- [10] H. Nguyen et al., "Animal recognition and identification with deep convolutional neural networks for automated wildlife monitoring," in *2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 2017, pp. 40–49. <https://doi.org/10.1109/DSAA.2017.31>.
- [11] A. Gomez, G. Diez, A. Salazar, and A. Diaz, "Animal identification in low quality camera-trap images using very deep convolutional neural networks and confidence thresholds," in *International Symposium on Visual Computing*, 2016, pp. 747–756. https://doi.org/10.1007/978-3-319-50835-1_67
- [12] J. Xiao, G. Liu, K. Wang, and Y. Si, "Cow identification in free-stall barns based on an improved Mask R-CNN and an SVM," *Comput. Electron. Agric.*, vol. 194, 2022, <https://doi.org/10.1016/j.compag.2022.106738>
- [13] S. Li et al., "Individual dairy cow identification based on lightweight convolutional neural network," *PLoS One*, vol. 16, no. 11, November, 2021, <https://doi.org/10.1371/journal.pone.0260510>
- [14] M. Hussain, "YOLO-v1 to YOLO-v8, the Rise of YOLO and Its Complementary Nature toward Digital Manufacturing and Industrial Defect Detection," *Machines*, vol. 11, no. 7, p. 677, 2023, <https://doi.org/10.3390/machines11070677>
- [15] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961–2969, <https://doi.org/10.1109/ICCV.2017.322>
- [16] S. A. Magalhães et al., "Evaluating the single-shot multibox detector and YOLO deep learning models for the detection of tomatoes in a greenhouse," *Sensors*, vol. 21, no. 10, p. 3569, 2021, <https://doi.org/10.3390/s21103569>
- [17] CVAT, "CVAT-Open Data Annotation Platform," 2023. <https://www.cvat.ai/>

BIOGRAFÍAS

Filio Carrasco Sauñe, bachiller en Ingeniería Informática y Sistemas de la Universidad Nacional Micaela Bastidas de Apurímac, actualmente es analista y programador de software de la empresa Avalon Soluciones S.A.C.

Manuel Jesús Ibarra Cabrera, doctor en ciencias de la computación e investigador en las áreas de ingeniería de software, serious game, informática educativa, computación móvil, IoT e industria y sociedad. Docente universitario de pre y pos grado en la Universidad Nacional Micaela Bastidas de Apurímac, Universidad Tecnológica de los Andes, Universidad Nacional del Altiplano y la Universidad Nacional Mayor de San Marcos. Actualmente, es miembro activo de conferencias importantes en latinoamerica: LACLO, CLEI, CONTIE, HCI, SBTIC, CISTI, Decisioning y otros.