

Monitoreo basado en visión computacional de aves en peligro de extinción de la reserva Pantanos de Villa

Luis Giancarlo Sánchez Huamán

20161333@aloe.ulima.edu.pe

Universidad de Lima, Perú

doi: <https://doi.org/10.26439/ciis2021.5640>

Los gobiernos, junto con numerosas organizaciones de conservación, trabajan para proteger las aves, en especial en las reservas naturales como los Pantanos de Villa, Perú. Sin embargo, sus poblaciones están descendiendo a escala global, fundamentalmente por causa de la contaminación. Por ello, es necesario realizar censos permanentes para monitorear la cantidad de aves, de manera ágil y a bajo costo, sin requerir la intervención de un experto humano. El presente trabajo propone un sistema de monitoreo formado por un modelo de detección utilizando Mask R-CNN con un *dataset* público Open Images Dataset V6 y un clasificador Deep Convolutional Neuronal Network con un *dataset* de cuatro especies de aves de la reserva de Pantanos de Villa. El sistema se basa en imágenes capturadas por los propios trabajadores, donde se realice la detección y clasificación. Tras varias evaluaciones diarias se determinará la población existente de aves y, cuando se presente una disminución, la reserva tomará acciones. El sistema propuesto logró detectar y clasificar las especies de aves para la realización del monitoreo con una exactitud de 82,49 % logrando mejorar los valores obtenidos previamente empleando únicamente el clasificador.

MONITORING BASED ON COMPUTER VISION OF ENDANGERED BIRDS IN THE PANTANOS DE VILLA RESERVE

Governments, along with numerous conservation organizations, work to protect birds primarily in nature reserves, specially in the Pantanos de Villa reserve in Peru. Therefore, it is necessary to carry out permanent censuses to monitor these populations, in an agile way and at a low cost, without requiring the intervention of a human expert. The present work proposes a monitoring system consisting of a detection model using Mask R-CNN with a public Open Images Dataset V6 dataset and a Deep Convolutional Neuronal Network classifier with a dataset of four bird species from the Pantanos de Villa reserve, Peru. The system will be used to monitor images captured by the workers themselves, where the detection and classification are carried out; After several daily evaluations, the existing population of the birds will be determined, and when there is a decrease, the reserve will take action. The proposed system managed to detect and classify bird species for monitoring, obtaining an accuracy of 82,49%, improving the values previously obtained using only the classifier.

Monitoreo basado en visión computacional de aves en peligro de extinción de la reserva Pantanos de Villa

Luis Giancarlo Sánchez Huamán
20161333@aloe.ulima.edu.pe

Resumen. Los gobiernos, junto a numerosas organizaciones de conservación, trabajan para proteger las aves, principalmente en las reservas naturales. Sus poblaciones están descendiendo a nivel global, inclusive numerosas especies están en peligro de extinción, fundamentalmente por efectos de la contaminación. Prueba de ello es que en el año 2012 se presentaron reportes de aves muertas en la reserva de Pantanos de Villa, Perú, por lo que es necesario realizar censos permanentes para monitorear dichas poblaciones, de manera ágil y a bajo costo, sin requerir la intervención de un experto humano. El presente trabajo propone un sistema de monitoreo formado por un modelo de detección utilizando Mask R-CNN con un *dataset* público Open Images Dataset V6 y un clasificador Deep Convolutional Neural Network con un *dataset* de cuatro especies de aves de la reserva de Pantanos de Villa, Perú; el sistema será empleado para realizar un monitoreo a partir de imágenes capturadas por los propios trabajadores, donde se realice la detección y clasificación. Tras varias evaluaciones diarias se determinará la población existente de las aves, y cuando se presente una disminución, la reserva tomará acciones. El sistema propuesto logró detectar y clasificar las especies de aves para la realización del monitoreo con una exactitud de 82,49 % logrando mejorar los valores obtenidos previamente empleando únicamente el clasificador.

Introducción

En el año 2009 el Perú se encontró en el tercer lugar del mundo en riqueza de especies de aves, con alrededor de 1729 especies. La mayoría de estas se encuentran en zonas protegidas por el país como el Refugio de Vida Silvestre los Pantanos de Villa (RVSPV).

En el 2012, el Gobierno peruano registró un gran número de aves marinas muertas en la costa peruana (Valle B., 2013), lo cual generó la realización de un monitoreo por parte del RVSPV que se volvió prioritario. Según Valle B. (2013), "se registraron 199 aves muertas, que corresponden a ocho especies: pelicano peruano, piquero, cormorán neotropical, cormorán guanay, gaviota peruana, gaviota de capucha gris, gaviota gris, gaviota de Franklin".

Adicionalmente, se han realizado otros estudios ornitológicos como fue el caso de Iannacone *et al.* (2010) que demandó una gran cantidad de tiempo y personas; por ello, la realización de estos estudios es complicada al no contar con un sistema que lo facilite. Igualmente en la actividad diaria, los trabajadores no cuentan con una herramienta de apoyo que permita dejar de lado el empleo de técnicas simples, pero que demandan una cantidad de tiempo considerable (Requena *et al.*, 2014).

Para apoyar en esa problemática, se propuso un sistema de monitoreo para las especies en peligro con la finalidad de tener un control de estas.

Materiales y métodos

Se construyó un sistema de monitoreo compuesto de Mask Region Based Convolutional Neural Network (Mask R-CNN) para la detección y Deep Convolutional Neural Network (DCNN) para la clasificación de las aves (véase la figura 1).

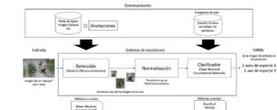


Figura 1. Diagrama de la metodología propuesta. Elaboración propia

Se utilizó un *dataset* público (Open Images Dataset V6) para la detección y otro privado para la clasificación. Para la detección se utilizaron 500 imágenes (300 de entrenamiento, 100 de validación y 100 de prueba) y para la clasificación 2000 imágenes (1200 de entrenamiento, 400 de validación y 400 de prueba), 500 por cada especie aproximadamente.

En la etapa de detección se empleó un modelo preentrenado del repositorio de Abdulla (2017) modificando la salida y entrenándolo con Open Images Dataset V6; por otro lado, en la etapa de clasificación se construyó un modelo *ad hoc* de 7 Convolutional Layer (ReLU function), 7 Pooling Layer, 7 Dropout Layer, 1 Flatten Layer y 1 Dense Layer (Softmax Function) utilizando *data augmentation*.

Se implementaron los modelos utilizando como lenguaje de programación Python y librerías de TensorFlow, Keras, NumPy, Scikit-learn y OpenCV en Google Colab.

Resultados

Se realizó una comparación entre un modelo de clasificación directa y el sistema de monitoreo (detección y clasificación conjunta), con la finalidad de conocer el impacto de emplear la etapa de pre-procesamiento. El mismo *dataset* ha sido utilizado en ambos para evaluar su rendimiento.

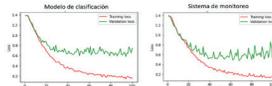


Figura 2. Pérdida por épocas para la etapa de entrenamiento y validación del modelo de clasificación y el sistema de monitoreo. Elaboración propia

Las gráficas presentadas en la figura 2 corresponden a la pérdida. Ambas tienen una tendencia a reducir el valor de la pérdida por cada época, siendo el modelo de clasificación más estable en el tiempo y el sistema de monitoreo menos estable, pero con mayor tendencia a obtener menores valores. Para el subconjunto de entrenamiento, el modelo de clasificación tiende a valores cercanos a 0,2 y para el subconjunto de validación, los valores oscilan entre 0,6 a 0,78; ambas líneas se mantienen cercanas hasta la época 18. Por otro lado, el sistema de monitoreo para el subconjunto de entrenamiento presenta valores cercanos a 0,13 y para el subconjunto de validación los valores oscilan entre 0,5 a 0,8, siendo inferiores a los presentados por el modelo de clasificación. Las líneas presentan valores cercanos hasta la época 24.

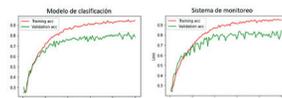


Figura 3. Exactitud por épocas para la etapa de entrenamiento y validación del modelo de clasificación y el sistema de monitoreo. Elaboración propia

En la figura 3 ambas gráficas corresponden a la exactitud y presentan un comportamiento similar al de las pérdidas, donde el modelo de clasificación presenta valores con poco margen de variación, mientras que el sistema de monitoreo tiene más margen entre los valores. Para el subconjunto de entrenamiento el modelo de clasificación presenta valores cercanos a 1, mientras que para el subconjunto de validación los valores se acerca a 0,8 por cada época, aunque existen ciertas épocas donde el valor baja hasta 0,76; cabe resaltar que ambas líneas crecen juntas hasta la época 18. Por otro lado, el subconjunto de entrenamiento para el sistema de monitoreo presenta valores cercanos a 1 mientras que el subconjunto de validación presenta valores cercanos a 0,84, pero en algunas épocas el valor baja hasta 0,75 presentando un mayor rango que el modelo de clasificación, pero obteniendo mejores valores dentro de ello; las líneas se mantienen unidas hasta la época 24.

Por otro lado, para el subconjunto de pruebas el sistema obtuvo una exactitud de 82,49 % y una pérdida de 0,51 siendo superior a las métricas presentadas solo por el clasificador, 80,62 % de exactitud y 0,64 de pérdida; esto demuestra un

impacto real al utilizar la etapa de detección previamente a la etapa de clasificación.

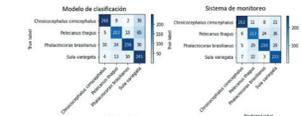


Figura 4. Matriz de confusión del modelo de clasificación y el sistema de monitoreo. Elaboración propia

En la figura 4 las predicciones para ambas matrices de confusión presentan valores similares, pero hay registros de aves para el sistema que presentan mejores resultados, como el caso de *Choricocephalus cirrocephalus* y *Pelecanus thagus*, y otro presenta peores resultados, *Sula variegata*. El acierto o fallo que presenta el sistema corresponde a las imágenes de entrada, ya que se ha eliminado el fondo obteniendo solo las características más importantes (efecto positivo), pero perdiendo calidad por el recorte y el dimensionamiento del ave (efecto negativo).

En la tabla 1 se presentan las métricas obtenidas para el modelo de clasificación y el sistema.

Etapas de clasificación	Exactitud	Pérdida	Precisión	Exhaustividad	Valor-F
Modelo de clasificación	80,62 %	0,64	80,74 %	82,37 %	80,77 %
Sistema de monitoreo	82,49 %	0,51	82,54 %	82,37 %	81,87 %

Tabla 1. Resultados de desempeño del modelo de clasificación y sistema de monitoreo. Elaboración propia

El sistema obtuvo un 82,49 % en el reconocimiento de la especie del ave presentada, cabe resaltar que la imagen puede presentar varias aves y la etapa de detección se encarga de separarlos para pasar a la etapa de clasificación.

Conclusiones

El presente trabajo tuvo como objetivo elaborar un sistema de monitoreo de especies de aves en peligro de una reserva natural basado en Visión Computacional; para ello, se creó el sistema a partir de la unión de dos modelos, detección (Mask R-CNN preentrenado) y clasificación (DCNN propio), empleando un *dataset* de 500 imágenes de Open Images Dataset V6 de Google para la detección y 2000 imágenes recopiladas de eBirds.org para la clasificación; obteniendo una exactitud de 82,49%.

La inclusión del modelo de detección como etapa anterior a la clasificación tuvo un impacto positivo permitiendo la mejora de la exactitud en un 1,87 %. En esta versión el sistema solo trabaja con cuatro especies de aves; sin embargo, esto puede expandirse trabajando con un *dataset* robusto que proporcione una mayor cantidad de especies de aves. El principal aporte de esta investigación es que permite la detección y clasificación de algunas aves en peligro de la investigación de Valle B. (2013), sin tener que recurrir a expertos ornitólogos; además facilita el monitoreo diario por parte de los propios trabajadores, ya que lo realizarán con imágenes y conocerán la población de cada especie. De este modo, si existe un caso de decrecimiento poblacional o desaparición de la especie, la reserva tendrá conocimiento para tomar acciones.

Referencias

Abdulla, W. (2017). *Mask R-CNN for Object Detection and Instance Segmentation on Keras and TensorFlow*. GitHub Repository.

Iannacone, J., Atasi, M., Bocanegra, T., Camacho, M., Montes, A., Santos, S., Zúñiga, H., y Alayo, M. 2010. Diversidad de aves en el humedal pantanos de villa, Lima, Perú: período 2004-2007. *Biota Neotropica*, 10(2), 295-304.

Kumar, A., y Das, S. D. (2018). Bird Species Classification Using Transfer Learning with Multistage Training. *Communications in Computer and Information Science*, 1019 CCIS (pp. 1-9).

Requena, F., Ascencios, M., Alva, J., y Pineda, A. (23 de enero del 2014). *Paraiso ecológico. El Peruano*.

Torres, J. 2018. *Deep Learning: Introducción práctica con Keras* (2.ª ed.). Creative Commons.

Valle B., D. F. (2013). Evaluación de los efectos del evento de mortandad de aves marinas durante mayo de 2012 sobre la población de aves en el Refugio de Vida Silvestre Los Pantanos de Villa, Lima-Perú. *Boletín Informativo de La Unión de Ornitólogos Del Perú*, 8(1), 44-58.

Agradecimientos

Quiero agradecer a mis profesores, Paul Quiroz, Victor Hugo Ayma, Rosario Guzmán y José Antonio Taquia, quienes con sus conocimientos me apoyaron a través de cada una de las etapas de esta investigación. Asimismo, agradecer a mi familia por el apoyo para terminar esta investigación.



UNIVERSIDAD
DE LIMA