

Optimizando la identificación y estudio biológico de aves mediante IA: resultados preliminares

Teodoro Ismael Fernandez Cruz¹, Octavio Ascagorta¹, Francisco Ramiro Iaconis³, Glenda Denise Hevia², María Débora Pollicelli^{1,2}

¹ Laboratorio de Investigación en Informática (LINVI), Departamento de Informática, Facultad de Ingeniería, Universidad Nacional de la Patagonia San Juan Bosco (UNPSJB), Puerto Madryn, Chubut, Argentina

² Centro para el Estudio de Sistemas Marinos, Consejo Nacional de Investigaciones Científicas y Técnicas (CESIMAR - CONICET), Puerto Madryn, Chubut, Argentina

³ Instituto de Física del Sur, Departamento de Física, Universidad Nacional del Sur (UNS-CONICET)

Abstract. La fotografía de fauna silvestre constituye una herramienta fundamental en la obtención de registros de individuos de aves anilladas (banderillas con códigos únicos) para el estudio de especies de amplia distribución geográfica. El desarrollo tecnológico, junto con el incremento de iniciativas de ciencia ciudadana, han contribuido a enriquecer estos registros de manera exponencial. Esta línea de investigación busca optimizar el procesamiento y el análisis de grandes volúmenes de imágenes con sus metadatos asociados, así como facilitar la exploración e interpretación de datos mediante herramientas analíticas espacio-temporales. Se destacan como resultados preliminares el desarrollo de modelos basados en deep learning para la foto-identificación automatizado del Chorlo Doble Collar (*Charadrius falklandicus*), un ave playera que típicamente reproduce en las zonas costeras de Patagonia. En particular, se exponen dos modelos obtenidos que demuestran la viabilidad de este enfoque para realizar una clasificación dicotómica de imágenes de presencia/ausencia de la especie y la posterior detección de la banderilla, que permite arribar a la identificación del individuo.

Keywords: aves · foto-identificación · visión por computadora

1 Introducción

El Chorlo Doble Collar (*Charadrius falklandicus*), en adelante CDC, es un ave playera neotropical del sur de Sudamérica, con amplia distribución en la Patagonia. Individuos de dicha especie han sido marcados durante campañas científicas en sitios reproductivos ubicados al noreste de la provincia de Chubut, Argentina. [1, 2]. El marcaje consistió en la colocación de una banderilla con un código alfanumérico en la tibia y un anillo metálico en el tarso siguiendo los lineamientos del Protocolo Panamericano de Marcaje para Aves Playeras del Hemisferio Occidental [5]. Posteriormente, numerosos individuos han sido reavistados en diferentes estaciones del año [1, 2]. Los registros de avistamientos realizados por

científicos y actores no científicos han resultado extremadamente útiles para apoyar la investigación y el monitoreo en campo de poblaciones de aves [8, 3] y contribuyen a mejorar enormemente los estudios de poblaciones de aves mediante la foto-identificación [8]. El uso de cámaras con lentes potentes constituye una herramienta eficaz para la detección de aves marcadas. Permite la captura precisa de los códigos de las banderillas mediante reavistamientos in situ y proporciona una valiosa contribución mediante la captura de fotografías, junto con metadatos geo-temporales. Esta información es utilizada en estudios biológicos como filopatria, distribución poblacional, patrones migratorios, uso de hábitat, impactos antrópicos, entre otros. Esta línea de investigación busca automatizar el procesamiento de grandes volúmenes de imágenes, acelerando la extracción, análisis y disposición de datos para el estudio biológico de la especie. La hipótesis fundamental de esta investigación subyace en que por un lado, el desarrollo de algoritmos de procesamiento de imágenes basados en visión por computadora permiten la detección de objetos de interés, de forma ágil y precisa [9]. A la vez el reconocimiento óptico de caracteres (OCR) permite la extracción de caracteres alfanuméricos [5, 7]. En conjunción se logra un procesamiento automatizado y masivo, en la detección del CDC y la banderilla, y la posterior extracción del código alfanumérico (flagID) de las mismas. Adicionalmente, los datos extraídos de la imagen, sus metadatos y otros registros de estudio son plausibles de exploración y visualización al desarrollar herramientas que permitan una analítica visual espacio-temporal y faciliten su interpretación [6].

2 Materiales y métodos

La metodología propuesta se desarrolla de la siguiente manera: (1) preparación del conjunto de datos, (2) entrenamiento y validación (incluyendo el testeo y ajustes) de los modelos de detección sobre imágenes, (3) extracción de caracteres mediante OCR, (4) desarrollo de prototipos de análisis y visualización de los datos y (5) ensamble e implementación del flujo de trabajo completo. El dataset utilizado se compone de un total de 795 imágenes, 450 tienen presencia de CDC, con y sin banderillas adosadas en las tibias (`bird_cdc`) y 345 imágenes con ausencia de CDC y presencia de otras especies de aves (`no_cdc`). Las imágenes provienen de estudios científicos, aportes de actores no científicos (ciencia ciudadana) y fotografías de dominio público, junto con sus metadatos. Luego, se establecen los objetos de interés (ROIs) a ser anotados sobre el dataset: (`bird_cdc`, `no_cdc` y `flagID`). Las anotaciones manuales se realizan mediante la herramienta LabelSudio. Finalmente se efectúa la revisión y validación por parte de expertos. Se proponen dos modelos basados en IA: (1) un clasificador de imágenes para filtrar imágenes con presencia-ausencia de CDC y (2) un detector del CDC y la banderilla (`flagID`) si hubiera. En el proceso de desarrollo de modelos el conjunto de datos es dividido en dos partes: entrenamiento (80%) y validación (20%). La división del conjunto de imágenes se realizó de manera estratificada. Se aplicó aumento de datos, incluyendo: flip horizontal, rotación y deformación, zoom, modificación de brillo y contraste, etc. Para cada método se configuraron

los parámetros adecuados de ajustes, entre estos parámetros se incluye la probabilidad de aplicar las transformaciones. El desarrollo de modelos fue hecho enteramente en Python, con bibliotecas de código abierto para aprendizaje profundo. Se considera el uso de las métricas típicamente utilizadas en problemas de clasificación y detección de objetos en imágenes: accuracy, precisión, recall, F1-score. Para medir el rendimiento en detección de objetos se utilizan dos métricas de Mean Average Precision (mAP): mAP@0.5, a un umbral de Intersection over Union (IoU) de 0.5 y mAP@0.5:0.95, con múltiples umbrales de IoU, desde 0.5 hasta 0.95 en incrementos de 0.05.

Modelo de clasificación de imágenes. Este modelo fue desarrollado con la librería Fastai [4]. Se utilizaron las imágenes clasificadas con las clases "bird_cdc" y "no_cdc", conformando el conjunto de entrenamiento con 504 y de validación con 132 imágenes. El modelo utiliza una arquitectura ResNet50, con pesos preentrenados. Para abordar un posible desequilibrio de clases, se emplea una función de pérdida ponderada. Las imágenes se preprocesan a un tamaño de 224x224 píxeles aplicando "Squish". Fue entrenado en 20 epochs y utiliza la función de pérdida CrossEntropy.

Modelo de detección de ROIs: individuo de CDC y banderilla. Este modelo fue diseñado para identificar la presencia y la ubicación de las ROIs, es decir de detectar individuos de CDC y banderillas (etiquetadas como flagID). Desarrollado mediante transfer learning (fine-tuning) basado en YOLO (You Only Look Once) V7 [10], con las imágenes del conjunto "bird_cdc" del dataset. El modelo utiliza la arquitectura E-ELAN con pesos pre-entrenados con COCO Dataset. Se aplicó el método letterbox, para el redimensionamiento de las imágenes a 640x640. Fue entrenado con 300 epochs, usando función de pérdida computeLossOTA que usa como base FocalLoss con algunos cambios.

3 Resultados y Discusión

El clasificador muestra un rendimiento con una precisión del 98.59%, una sensibilidad (recall) del 99.99%, una precisión del 98.5% y un F1 del 99.29%. En el caso del modelo de detección de ROIs el modelo presentó una precisión del 92.14%, sensibilidad del 88.15%, mAP@.5 del 89.38% y mAP@.5:.95 del 68.02%. La precisión del 98.59% del clasificador indica que el modelo presenta un bajo número de falsos positivos. El valor de sensibilidad obtenido, muestra la capacidad de detectar correctamente casi todos los casos de la clase objetivo. El F1 Score de 99.29% refleja un buen equilibrio entre la precisión y la sensibilidad. Estos resultados sugieren que el clasificador es altamente efectivo en la clasificación de imágenes con presencia de CDC. De esta forma, se descartan imágenes que no sean útiles, comúnmente otras especies de aves o capturas fotográficas fallidas. La precisión, la sensibilidad y el mAP@.5 del modelo de detección de ROIs son altas, lo que indica que el modelo es preciso y capaz de identificar correctamente la mayoría de las ROIs con un valor significativo de IoU. Sin embargo, un valor más bajo en mAP@.5:.95 muestra que la precisión del modelo disminuye a medida que aumenta el IoU.

4 Conclusión y trabajos futuros

Los resultados preliminares presentados, demuestran la capacidad de clasificación de imágenes con presencia de CDC, y la posterior detección y segmentación automática de las ROIs: CDC y banderilla. Esto evidencia la factibilidad de brindar una solución eficiente con la incorporación de IA al flujo de trabajo de los estudios biológicos. Al reducir la carga de tiempo y esfuerzo cognitivo en tareas que actualmente se realizan de forma manual, permitirá gestionar de manera eficiente el volumen de datos. Se espera desarrollar analíticas visuales que involucren tanto los datos extraídos de las imágenes, sus metadatos, así como otros registros de interés. Los resultados esperados serán un aporte no solamente para las investigaciones biológicas y la articulación de éstas con la ciencia ciudadana, sino también como contribución metodológica en el campo de procesamiento de imágenes y datos provenientes de diversas fuentes.

References

1. Hevia, G.D., Dallas Jordan, F., Terorde, N., Bertellotti, N.M., D'amico, V.L.: Resightings of two-banded plovers (*Charadrius falklandicus*) during the breeding season in coastal chubut, patagonia, argentina. *Ornitología Neotropical* (2018)
2. Hevia, G.D., D'Amico, V.L., Bertellotti, M.: Efectos de las actividades humanas sobre los parámetros reproductivos y la condición física de chorlos doble collar (*Charadrius falklandicus*) en chubut, patagonia. *El Hornero* **35**(2), 148–148 (2020)
3. Hevia, G.D., Fiorito, C.D., Terorde, N., Jordan, F.D., Podestá, D.H., Esmoris, A., Villareal, R.F., Burgueño, L., Corbalán, N.S., Chaparro, M.A., et al.: Improving resightings of two-banded plovers flag codes by using digital photo-identification in patagonia, argentina (2020)
4. Howard, J., Gugger, S.: Deep learning for coders with fastai and pytorch: Ai applications without a phd; o'relly media. Inc.: Sebastopol, CA, USA (2020)
5. Howes, L.A., Béraud, S., Drolet-Gratton, V.: Pan american shorebird program shorebird marking protocol. <https://www.shorebirdplan.org/wp-content/uploads/2016/08/PASP-Marking-Protocol-April-2016.pdf> (2016)
6. Mahendiran, M., Parthiban, M., Azeez, P.A., Nagarajan, R.: In situ measurements of animal morphological features: A non-invasive method. *Methods in Ecology and Evolution* **9**(3), 613–623 (2018)
7. Santangeli, A., Chen, Y., Boorman, M., Sales Ligeró, S., Albert García, G.: Semi-automated detection of tagged animals from camera trap images using artificial intelligence. *Ibis* **164**(4), 1123–1131 (2022)
8. Sullivan, B.L., Wood, C.L., Iliff, M.J., Bonney, R.E., Fink, D., Kelling, S.: ebird: A citizen-based bird observation network in the biological sciences. *Biological conservation* **142**(10), 2282–2292 (2009)
9. Vigneshwaran, P., Ratnarajah, N.: Detection of wildlife animals using deep learning approaches: A systematic review. 2021 21st International Conference on Advances in ICT for Emerging Regions (ICTer) pp. 153–158 (2021), <https://api.semanticscholar.org/CorpusID:248851469>
10. Wang, C.Y., Bochkovskiy, A., Liao, H.Y.M.: Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 7464–7475 (2023)