

AUTORES:

Sieiro, R.¹
Mtnéz. de Pisón, F. J.¹
González, F.²



¹Grupo EDMANS
www.mineriadatos.com
Universidad de La Rioja
c/ Luis de Ulloa 20
26004 Logroño



UNIVERSIDAD
DE LA RIOJA



MUSEO DE LA
NATURALEZA
DE CANTABRIA

²Museo de la Naturaleza
de Cantabria
www.museosdecantabria.es

Bo. Carrejo, s/n, 39592
Cabezón de la Sal,
Cantabria, España



GOBIERNO
de CANTABRIA
VICERRECTORIA
COMISIVIA DE INVESTIGACIONES, IGUALDAD,
CULTURA Y DEPORTE

CONTACTO:

Félix González
fei2.ctv@gmail.com

1. Introducción

La utilización generalizada de grabadoras de ultrasonidos autónomas facilita la obtención de grandes cantidades de muestras. El procesado de las mismas se inicia mediante un filtrado que discrimina aquellos archivos que contienen señales realmente emitidas por murciélagos de los archivos que son considerados como ruido. La mayor o menor sensibilidad para su detección puede complicar de forma importante posteriores análisis y distorsionar los resultados, ya que un elevado número de falsos positivos obliga a revisiones detalladas no siempre posibles, en tanto una selección muy conservadora de las muestras puede omitir una parte importante de la información.

El objetivo de este trabajo se centró en el desarrollo de modelos de aprendizaje profundo (*Deep Learning*, DL) capaces de discriminar con alta precisión las grabaciones de murciélagos de aquellas que solo correspondían a ruido ambiental (insectos, lluvia, viento...).

2. Metodología

A partir de 152 606 archivos de 0.1 s a 5 s de duración grabados en 12 localidades del área cántabra (figura 1) se seleccionaron 7751 muestras de ruido y murciélagos que presentaron alta dificultad en el proceso de clasificación. Los casos se etiquetaron de la siguiente forma: *Murciélagos* (55,45%), *Ruido* (27,88%) y *Ruido+Murciélagos* (16,67%). En esta última clase se incluyeron aquellas muestras con alto nivel de ruido donde también aparecían una o varias especies de murciélagos.

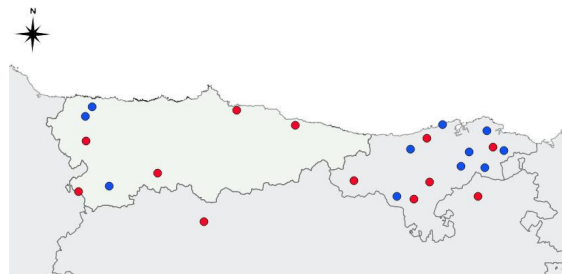


Figura 1. Distribución de las localidades de procedencia de las grabaciones de sonido utilizadas (localidades de testeo en azul).

Inicialmente, se eligieron las arquitecturas de DL más modernas que previamente ya habían sido entrenadas para la clasificación de objetos con la base de datos *Imagenet*, formada por 14 millones de imágenes. Con la nueva base de datos se reentrenaron los modelos para especializarlos en la tarea de identificar señales de murciélagos a partir de los espectrogramas de los audios.

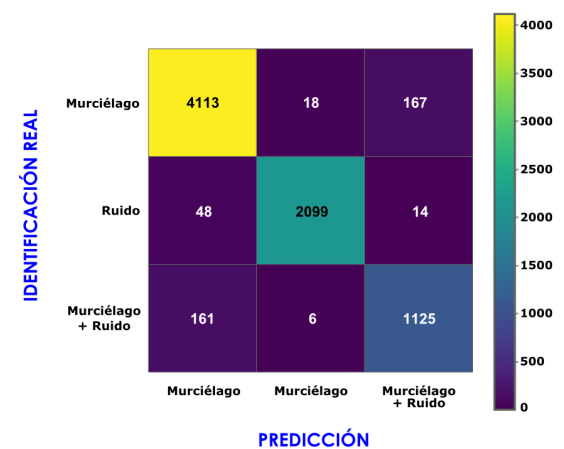
Se realizaron más de 150 experimentos con diversas arquitecturas de redes neuronales y combinaciones de hiperparámetros relacionados con el entrenamiento de los mismos y la creación de los sonogramas. El entrenamiento de los modelos se realizó con la función de pérdida *CrossEntropyLoss* y la validación/selección de los modelos se basó principalmente en la métrica *F1-Score Macro*. El ajuste de los parámetros se realizó en varias etapas comenzando por los parámetros relacionados con la creación de los sonogramas (número de ventanas FFT, máxima y mínima frecuencia, etc.), siguiendo por la selección de los métodos de aumento de datos (*augmentation*) más adecuados tanto en audio (inclusión de diversos tipos de ruido, modificación de ganancia, distorsión de la señal, etc.) como en imagen (deformación en tiempo, enmascaramiento en tiempo y frecuencia, etc.), y finalizando con la selección de las mejores arquitecturas de DL y el ajuste de sus hiperparámetros (parámetros de optimización y entrenamiento, arquitectura de la red, umbrales de decisión, etc.).

	MURCIÉLAGOS	RUIDO	MURCIÉLAGOS + RUIDO
5 Folds_LOSS=0.4893			
5 Folds_ACC=0.9689			
5 Folds_F1Score_Murc=0.9922			
5 Folds_AUC=0.9905			
Precision	0.9516	0.9887	0.8614
Recall	0.9565	0.9713	0.8707
F1Score	0.9540	0.9799	0.8661
Num	4298	2161	1292

Tabla 1. Métricas de validación cruzada con 5 pliegues obtenidas con uno de los mejores modelos, *EfficientNet-B0*.

De entre ellos, se eligieron los cinco mejores modelos cuya combinación de predicciones alcanzó una precisión del 97,3 % en una muestra de validación de 5000 archivos difíciles de clasificar correctamente por programas de identificación de especies y seleccionados de entre 380 725 archivos grabados en otras 12 localidades de la misma área.

Figura 2. Matriz de confusión de un modelo *EfficientNet-B0*, obtenida con una validación cruzada de 5 pliegues.



3. Resultados

Los resultados mostrados en el cuadro corresponden a la combinación de los cinco modelos de DL (DL5_83) frente a tres herramientas comerciales (Sonochiro, SCH; Sonobat, SBT y Kaleidoscope, KCP) utilizando 500 muestras difíciles de clasificar por un experto.

DL5_83			SCH		
	BATS	NOISE		BATS	NOISE
precision	1	0.9393	precision	0.9922	0.9504
recall	0.944	1	recall	0.9552	0.9914
f1score	0.971	0.969	f1score	0.973	0.970
SBT			KCP		
	PRED_NOISE	PRED_BATS		PRED_NOISE	PRED_BATS
TRUE_NOISE	232	0	TRUE_NOISE	230	2
TRUE_BATS	15	253	TRUE_BATS	12	256
	BATS	NOISE		BATS	NOISE
precision	0.596	0.9393	precision	0.944	0.875
recall	0.556	1	recall	0.884	0.940
f1score	0.575	0.969	f1score	0.913	0.906
	PRED_NOISE	PRED_BATS		PRED_NOISE	PRED_BATS
TRUE_NOISE	131	101	TRUE_NOISE	218	14
TRUE_BATS	119	149	TRUE_BATS	31	237

La herramienta comercial que obtuvo mejores resultados fue SCH, por lo que se eligió para compararla directamente con la nueva propuesta (DL5_83). La herramienta SCH falló en 14 muestras y los modelos de DL en 15. Sin embargo, los modelos de DL no mostraron falsos positivos obteniendo una precisión del 100% para la clase *Murciélagos*.

Con respecto al tiempo de cálculo, SCH tardó en realizar un análisis completo 59 minutos mientras que los modelos DL redujeron a la mitad el tiempo de inferencia en la misma máquina.

4. Conclusiones

Estos resultados mejoran notablemente la capacidad de filtrado, eficiencia y sensibilidad de algunos programas utilizados frecuentemente en España para esta labor y hacen más precisa la determinación de la actividad en situaciones en las que es importante cuantificar con precisión la actividad real de las especies (p. e., parques eólicos).

5. Referencias

- Bertran, M., R. M. Alsina-Pagès & E. Tena, 2019: *Pipistrellus pipistrellus* and *Pipistrellus pygmaeus* in the Iberian Peninsula: An Annotated Segmented Dataset and a Proof of Concept of a Classifier in a Real Environment. *Appl. Sci.* 2019, 9, 3467; doi:10.3390/app9173467.
- Lemen, C. P., Freeman, J. White y B. Andersen, 2015: The Problem of Low Agreement among Automated Identification Programs for Acoustical Surveys of Bats. *Western North American Naturalist*, 75. doi: 10.3398/064.075.0210.
- Mac Aodha, O., R. Gibb, K. E. Barlow, E. Browning, M. Firman, R. Freeman, B. Harder, L. Kinsey, G. R. Mead, S. E. Newson, Ivan Pandourski, S. Parsons, J. Russ, A. Szodoray-Paradi, F. Szodoray-Paradi, E. Tilova, M. Girolami, G. Brostow & K. E. Jones, 2018: Bat detective-Deep learning tools for bat acoustic signal detection. *PLoS Computational Biology*, 14 (3): e1005995. <https://doi.org/10.1371/journal.pcbi.1005995>.
- Pytorch Image Models (timm). Github Repository, 2021: <https://github.com/fastai/timm-docs>. [consulta 31/mayo/2021].
- Xing C., J. Zhaob, Y. Chena, W. Zhou & A. C. Hughes, 2019: Automatic standardized processing and identification of tropical bat calls using deep learning approaches. *Biological Conservation*, vol. 241. <https://doi.org/10.1016/j.biocon.2019.108269>.