

# Reconocimiento de trinos de aves mediante redes neuronales artificiales: Un Enfoque Prototipo

Victoria V. Ludeña, Estudiante<sup>1</sup>, Ricardo J. Jiménez, Estudiante<sup>2</sup>, and Cristian I. Pinzón, Doctorado<sup>3</sup>

<sup>1,2</sup>Universidad Tecnológica de Panamá, Panamá, victoria.ludena@utp.ac.pa, ricardo.jimenez3@utp.ac.pa

<sup>3</sup>Universidad tecnológica de Panamá, Panamá, cristian.pinzon@utp.ac.pa

*Abstract– Panamá is a country rich in fauna, specifically has a variety of native bird species and is an obligatory migration bridge. However, illegal hunting and indiscriminate logging threaten the survival of some bird species. This paper presented a prototype of a bird recognition using convolutional neural networks (CNN), through the capture of trills and their conversion to a spectrogram image to perform the classification and later identification using an application based on an expert system. The sections described the software tools and artificial intelligence algorithms used and presented a conceptual model of the prototype built. For the training of the CNN, a data set was developed based on ten birds from the region studied. Finally, the results obtained from the tests conducted on the prototype were presented, where an accuracy rate of approximately 90% was achieved. In a next stage, its application in primary schools as an innovative tool for learning outside the classroom is envisaged.*

## I. INTRODUCCIÓN

Panamá es un país tropical rico en biodiversidad, con una fauna variada. En cuanto a aves se refiere, es un país inclusive siendo pequeño en extensión territorial, según datos publicados por Audubon [1], cuenta con 1002 especies de aves reportadas de las cuales 177 son migratorias. Toda esta diversidad de aves comunes y migratorias en Panamá se requiere investigar, pues son un importante indicador de la salud de los ecosistemas naturales, tienen una función crucial en el control de plagas de los cultivos, en la dispersión de semillas en los hábitats naturales, inclusive en el desarrollo del turismo.

La identificación de las aves es un trabajo complicado, requiere de conocimiento y experiencia. Los expertos en el campo de la ornitología [2], requieren de años de experiencia para la tarea de identificar y clasificar las aves en las zonas geográficas donde viven y realizan migración. Aunque existen biólogos, ecologistas y ornitólogos con experiencia, reconocen que no es fácil llevar a cabo la tarea de identificación.

Aun con la gran diversidad de aves en Panamá, es importante mencionar que existen especies en peligro de extinción ya sea por el impacto del desarrollo agrícola y urbano en los bosques o por el tráfico ilegal de aves. Todo esto requiere de soluciones que contribuyan a proteger esta riqueza de nuestro país. Existen organizaciones encargadas de estudiar y dar a conocer la avifauna como lo son MiAmbiente [3], Audubon, Panamá Rainforest Discovery Center [4] y Ancón [5]. Sin embargo, se requieren más esfuerzos para proteger especies de aves en peligro de extinción y evitar que nuevas especies entre en esta lista.

**Digital Object Identifier:** (only for full papers, inserted by LACCEI).

**ISSN, ISBN:** (to be inserted by LACCEI).

**DO NOT REMOVE**

Con el objetivo de apoyar el trabajo de investigadores en el área de ornitología y la educación ambiental en las escuelas se propone facilitar la tarea de identificar ejemplares de aves mediante el procesamiento digital de señales sonoras, visión artificial, y técnicas de Deep Learning [6]; se plantea un prototipo de reconocimiento de sonidos de aves de la región, aplicando técnicas de inteligencia artificial.

El artículo se encuentra estructurado de la siguiente manera: la primera sección corresponde a la introducción, la segunda trata la descripción de la problemática que se plantea resolver, la tercera sección da a conocer la metodología de investigación empleada, en el cuarto apartado se describen los materiales y métodos utilizados para la construcción del prototipo, la quinta sección muestra los resultados, la última sección presenta las conclusiones del trabajo.

## II. ANTECEDENTES

### A. Trabajos de referencia

Se han identificado algunas soluciones existentes con el objetivo de ayudar a la identificación de aves. En la tabla I, se resumen los enfoques revisados en este artículo.

TABLA I  
COMPARACIÓN DE ENFOQUES DE IDENTIFICACIÓN DE AVES

DCNN [7]	Disimilitud [8]	Componentes principales [9]
Uso de tres tipos de espectrogramas, calculados a través de la transformada de Fourier de tiempo corto, la transformada de Cespstrum de frecuencia Mel y la transformada de Chirplet	Uso de la técnica de obtención de las propiedades de las imágenes del espectrograma, y el trabajo de disimilitud.	Análisis de los componentes del espectrograma en los cantos, haciendo uso de la técnica de Eigenfaces.

En base a estos trabajos, se plantea una solución innovadora utilizando algoritmos de Inteligencia Artificial, concretamente algoritmos de aprendizaje automático para identificar aves en regiones de Panamá.

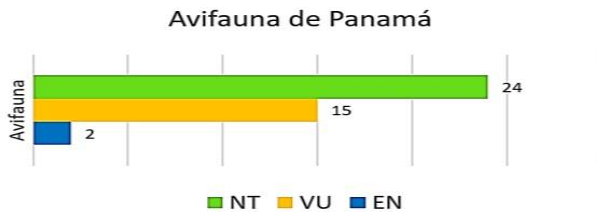
En esta etapa del proyecto, se ha delimitado un caso de estudio basado en la identificación de catorce aves frecuentes en la región destacando algunas como, el Sotorrey común (Troglodytes aedon) que se caracteriza por ser un ave pequeña de color café, el Bienteveo grande (Pitangus sulphuratus) que es un ave de color amarillo con café, y la Tangara azuleja (Thraupis episcopus), un ave de color celeste con pico negro, ver Fig. 1. Estas son especies nativas de Panamá.



Fig 1. Aves en estudio. a) Sotorrey común b) Bienteveo grande c) Tangara azuleja. Fuente: Audubon.

### B. Aves en Panamá

Debido a la posición geográfica de Panamá, la diversidad de aves es notoria y muy importante para el equilibrio de los ecosistemas locales. Hay especies propias de una región en específico, y enfrentan amenazas derivadas de las actividades del ser humano, principalmente contra su hábitat. En la Fig. 2 se muestra la situación de las aves amenazadas en Panamá.



NT: Casi amenazado; VU: Vulnerable; EN: En peligro

Fig. 2. Situación de las especies de aves amenazadas en Panamá. Fuente de datos: CBD Fourth National Report.[10]

Existen aproximadamente 1010 especies de aves en Panamá [11]. “De acuerdo con la Resolución No. DM-0657-2016 se declaran como amenazadas y/o en peligro de extinción alrededor de 30% de las especies del país”. El porcentaje puede ir en aumento si no se tiene conocimiento de las aves que viven en nuestro entorno para su protección debido a la cacería furtiva y tala indiscriminada del hábitat. Es importante mencionar que 12 especies son endémicas de Panamá.

¿Cuál es el problema por resolver? Por la gran variedad de aves con que cuenta Panamá, es importante que las personas pueden reconocer las aves de su entorno, y entender su importancia en el hábitat donde viven. Como se planteó anteriormente, la importancia de las aves se debe a que son unos de los indicadores de la salud de los ecosistemas naturales, adicional, su enorme importancia en la polinización y control de plagas en los cultivos. No educar a la población, y a las próximas generaciones, contribuirá a que nuestra riqueza en aves se siga perdiendo a un ritmo muy acelerado agravado ahora con la crisis climática.

Un enfoque educativo es salir y ver las aves en su entorno, estudiar sus comportamientos, y disfrutar sus cantos, sin embargo, no siempre se pueden observar fácilmente, y más si están en peligro de extinción con un reducido número de ejemplares. En este sentido, una pregunta lógica es como reconocer y estudiar las aves si no se pueden ver e identificar a simple vista.

Una solución posible es la identificación por el canto o trino de las aves. Este sonido se puede escuchar a muchos

metros y en algunos casos a kilómetros de distancia dependiendo del entorno. Considerando esta ventaja, se pueden construir herramientas que permitan capturar el trino a varios metros de distancia e identificar el tipo de ave que está en el entorno. En este punto se puede utilizar los potentes algoritmos que provee la inteligencia artificial para el reconocimiento de patrones.

Considerando todo esto, se plantea una pregunta de investigación. ¿Cuál es la arquitectura de software, basada en algoritmos de aprendizaje automático, para construir una solución que permita la identificación de aves a través del trino?

En este artículo se propone la construcción de un prototipo basados en aprendizaje automático como una respuesta a la pregunta de investigación.

### B. Materiales: Tecnologías implicadas

Las tecnologías empleadas y adaptadas al sistema se mencionan a continuación:

- Python [12]: La herramienta es utilizada para el desarrollo de los algoritmos de creación y obtención de características de un espectrograma, y prueba una red neuronal artificial.
- JavaScript [13]: Esta herramienta se requiere para programar la captura de audios y la interacción con los elementos HTML y CSS.
- Algunos paquetes, librerías y extensiones necesarias para el desarrollo del prototipo:
- Tensorflow [14]: Esta interfaz es indispensable en el proyecto ya que provee los algoritmos de aprendizaje automático necesarios dentro del proyecto.
- Librosa [15]: Utilizado principalmente para el análisis de sonido. El paquete es necesario para conseguir los espectrogramas, y características de los trinos de las aves.
- Numpy [16]: Se utiliza para representar valores del espectrograma, entre otras operaciones matemáticas.
- Flask [17]: En nuestro proyecto esta herramienta permite el uso de Python en un entorno web, para el manejo de solicitudes entre los lenguajes de programación.

## III. METODOLOGÍA

Se plantea una investigación aplicada, con un enfoque descriptivo, dirigida a solucionar un problema real. Se plantea con un enfoque descriptivo en esta primera etapa ya que se describe el problema, se recopila información y se plantea una posible solución.

A continuación, se plantea las metodologías utilizadas en esta propuesta:

### A. Metodología de investigación

#### a) Identificación del problema

Se realizó una búsqueda de herramientas locales para la identificación de aves con muy pocos resultados. Las herramientas encontradas tenían un carácter macro para identificar aves en muchos países, pero muy generalizado, y no enfocado a Panamá y sus regiones.

*b) Revisión bibliográfica*

Se efectuaron búsquedas de información científicas del tema. Como herramientas de búsqueda se utilizó el repositorio UTP-Ridda2 [18], Google Scholar [19] y otras fuentes de información.

*c) Recolección de datos*

Se hicieron grabaciones de aves frecuentes en la provincia de Veraguas, específicamente en los alrededores del distrito de Atalaya, ver Fig. 3. Esto se lleva a cabo en esta zona, debido a que el proyecto está siendo enfocado en esta región, además se complementaron con audios descargados mediante internet de las aves seleccionadas, los cuales permitieron obtener un registro de mejor calidad del canto.



Fig. 3. Área geográfica objeto de estudio. Fuente: Propia

*d) Diseño del modelo*

Se realizan varios diseños conceptuales del prototipo, y se escoge el diseño adecuado al problema. Luego de este paso se inicia la construcción y prueba del prototipo.

**B. Metodología de software**

Se plantea una aplicación de software desde la perspectiva de un sistema experto, para esto se requiere un marco estructurado de sistemas basados en el conocimiento. En este aspecto se plantea la utilización de la metodología CommonKADS [20].

**IV. PROTOTIPO PROPUESTO**

A continuación, se presenta el diseño conceptual en la Fig. 4 del prototipo planteado donde se muestran las herramientas de software y hardware a utilizar, las librerías y el flujo de datos que se debe realizar para el funcionamiento del prototipo.

A continuación, se explica el funcionamiento del prototipo paso a paso:

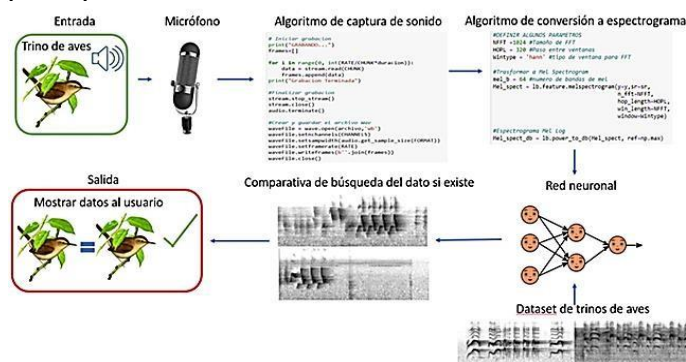


Fig. 4. Modelo conceptual del proyecto. Fuente: Propia

**A. Captación de audio**

El proceso comienza mediante la captura del trino del ave a través de un micrófono, y el empleo de un algoritmo de recolección de sonido mono canal. Se realizaron las grabaciones de poca duración para la creación del espectrograma, estas grabaciones fueron almacenadas en el formato “WAV”. En la Fig. 5 se representa la forma del audio captado.

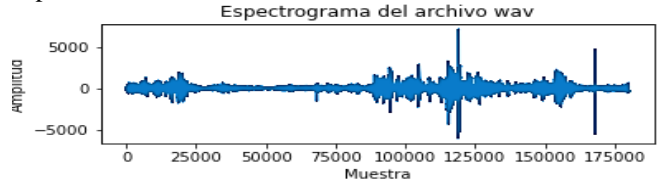


Fig 5. Ejemplo del espectrograma del archivo WAV. Fuente: Propia

**B. Creación del espectrograma**

Luego de realizar la captura del trino, del audio grabado se obtuvo la señal de entrada y la frecuencia de muestreo, posteriormente se calculó el espectrograma de Mel [21], más adelante esta salida es convertida en una imagen en escala de grises invertida, como se muestra en la Fig. 6. Para realizar la transformación, se limitaron los datos de la imagen entre 0 y 255, posteriormente se invirtió la escala de grises, restando a 255 el valor que poseen actualmente.

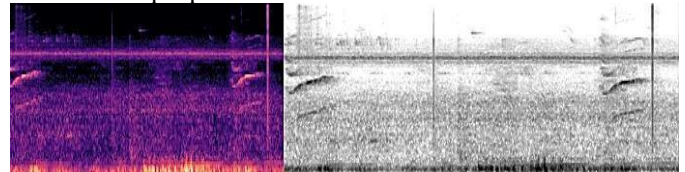


Fig 6. Conversión de RGB a una escala de grises. Fuente: Propia

**C. Red neuronal convolucional**

Se utilizó una red neuronal convolucional [22] (CNN) para encontrar patrones en la imagen del espectrograma. Esta red neuronal artificial realiza el proceso de identificación del ave con gran precisión.

**D. Conjunto de datos**

El conjunto de datos o dataset, es aquel medio por el cual se almacenaron datos selectos previamente recopilados. Los datos obtenidos en el caso presentado, se guardaron las imágenes del espectrograma de Mel de los trinos de las aves para el entrenamiento de la CNN.

**E. Desarrollo del prototipo**

En los siguientes párrafos se explica cómo se desarrolló el prototipo y un resumen de los entrenamientos llevados a cabo por la red CNN para realizar las pruebas del prototipo.

*a) Creación del prototipo*

El modelo escogido de la CNN para la creación y entrenamiento de la red neuronal fue de tipo Secuencial, compuesto por capas convolucionales, capas de agrupación máxima, una capa plana y capas densas.

Se creó una interfaz usando el lenguaje de programación JavaScript, que consta de 2 botones para grabar el trino del ave, y otro para identificar el ave grabada como se muestra en la Fig. 7.

## Identificación de aves



Fig 7. Interfaz del prototipo. Fuente: Propia

El primer botón realiza la función de grabar, en esencia se captura un audio de 10 segundos en un solo canal y se guarda en un archivo WAV.

El segundo botón ejerce la función de identificar el ave. El paso consiste en crear el espectrograma de Mel con el audio capturado y luego el modelo realiza la predicción, esta fue desarrollada por medio de la CNN entrenada, donde el modelo y los pesos se convierte en archivos .h5.

El siguiente paso consistió en cargar los archivos de audio, y luego se recibe el espectrograma como entrada, se clasifica en base a las categorías de aves que conoce la CNN.

## V. RESULTADOS

A continuación, se presentan los resultados obtenidos de las pruebas realizadas con el prototipo, ver la tabla II.

TABLA II  
PARÁMETROS DE LAS PRUEBAS DE ENTRENAMIENTO DE LA RED NEURONAL

#	Cantidad de aves	Tamaño data set	Épocas	Tasa de predicción	Tamaño de imagen
1	3	49	7	0.5897	100x100
2	3	131	20	0.9307	500x128
3	3	131	100	0.9901	500x128
4	3	131	35	0.9192	500x128
5	3	131	50	0.9596	250x250
6	8	888	75	0.9685	220x1540
7	13	1300	30	0.9685	224x224
8	13	1300	20	0.9556	224x224
9	14	1960	30	0.9118	128x128
10	14	1960	50	0.8889	128x128

Los resultados obtenidos fueron satisfactorios en esta etapa del prototipo. En el entrenamiento de la red neuronal se mostró un porcentaje de 88%, con la adición de nuevos conjuntos de aves, cada uno de estos conjuntos almacenaba 140 archivos de espectrogramas.

El entrenamiento se realizó bajo las condiciones de 50 épocas, se estructuró con tres capas de convolución, una capa plana para convertir la estructura dimensional en datos, además se hizo uso de técnicas de regulación con capas dropout a fin de evitar el sobre ajuste del modelo.

En algunos casos como tener dos cantos de aves mezclados o ruido ambiental se prioriza aquella opción que tenga una mayor probabilidad asociada, sin embargo, no siempre representará la opción definitiva o correcta.

### A. Mejoras futuras del prototipo

El prototipo se encuentra en desarrollo para mejorar los resultados obtenidos. En esta dirección se plantean algunas mejoras:

- Mejorar la interfaz gráfica del prototipo para tener una mejora visual.
- Aumentar el conjunto de datos de cada ave, además de agregar nuevas aves.
- Desarrollar una aplicación para dispositivos móviles accesible y de fácil uso para el usuario.

### B. Resultados de la encuesta

Se aplicó un análisis de percepciones a docentes de escuelas primarias, utilizando un muestreo por conveniencia con el propósito de entender la perspectiva de los docentes ante la implementación de herramientas tecnológicas en un ámbito educativo.

Los resultados preliminares obtenidos a partir de la encuesta muestran aceptación hacia el uso de recursos tecnológicos con un fin educativo, ver Fig. 8.

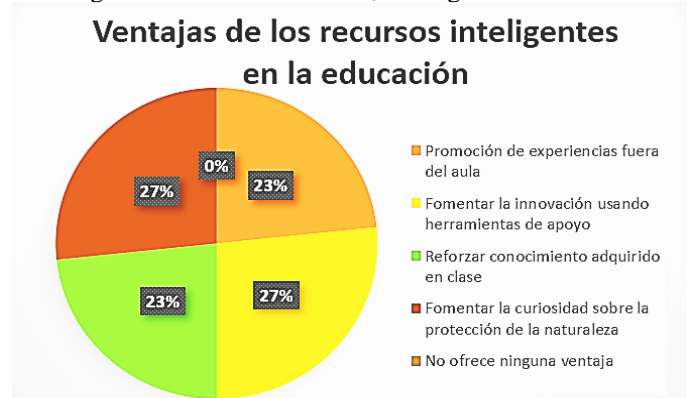


Fig 8. Ventajas de los recursos inteligentes en la educación. Fuente: Propia

A partir de los datos obtenidos podemos inferir que existe un interés en cuando a fomentar el uso de estas aplicaciones de apoyo, además que puede llegar a despertar la curiosidad de los estudiantes sobre la protección de la naturaleza.

Entre los resultados recopilados, se pudo distinguir otro dato relevante, el cual nos señala las posibles implementaciones que podría llegar a tener un software de este tipo, ver Fig. 9.

## Uso de software inteligente para identificar aves



Fig 9. Uso de software inteligente para reconocer aves. Fuente: Propia

### C. Impacto del proyecto

El proyecto impactará de forma positiva en áreas de investigación como apoyo a los especialistas. A nivel escolar como herramienta de complemento para actividades fuera del aula de clases y crear conciencia ambiental en los estudiantes. En el área de turismo como herramienta para la observación de aves.

## VI. CONCLUSIONES

Se resuelve la pregunta de investigación al construir un prototipo para reconocer las aves usando el trino, identificando las tecnologías de software y algoritmos para su construcción. Se puede destacar los siguientes aspectos:

- No es necesario un dispositivo especializado para captar el canto del ave, debido a que el prototipo funcionó con grabaciones obtenidas a través de dispositivos móviles de gama media a baja.
- El dispositivo pudo diferenciar el canto de las aves en presencia de poco ruido ambiental y a distancias aproximadas de hasta 5 metros.
- Según los datos obtenidos de la tabla II, la tasa de éxito en la precisión para identificar las aves mejora en función de la cantidad de datos disponibles en el conjunto de datos. En este sentido se trabaja para mejorar el prototipo aumentando la cantidad y calidad del dataset.
- Con las mejoras apropiadas esta herramienta puede ser de utilidad a investigadores, ornitólogos, biólogos, ecologistas, autoridades ambientales en su tarea de identificación de aves.
- A nivel escolar, en las escuelas primarias, se puede implementar como una herramienta de laboratorio fuera del aula de clases para enseñar a los alumnos a conocer la riqueza de su entorno y crear conciencia en la protección de las aves dada su importancia en los ecosistemas naturales.

## AGRADECIMIENTOS

Al Profesor Pablo Abrego, Dr. Carlos Bracho, Dr. Edmanuel Cruz, Doctores Elsy Abreu y Tito Ludeña por el apoyo en la realización de esta investigación.

## REFERENCES

- [1] “Inicio - Audubon Panamá.” <https://www.audubonpanama.org/>.
- [2] Avian Report, “Cómo Identificar Aves,” <https://es.avianreport.com/como-identificar-aves/>, Mar. 09, 2018.
- [3] “INICIO - MiAmbiente.” <https://www.miambiente.gob.pa/> (accessed Jun. 09, 2023).
- [4] “Birdwatchers’ favourite! - Panama Rainforest Discovery Center - Birdwatching Tower.” <https://pipelineroad.org/en/> (accessed Jun. 09, 2023).
- [5] “Home - ANCON Panamá.” <https://ancon.org/> (accessed Jun. 09, 2023).
- [6] “What is Deep Learning? | IBM.” <https://www.ibm.com/topics/deep-learning> (accessed Jun. 18, 2023).
- [7] F. Zhang, L. Zhang, H. Chen, and J. Xie, “Bird species identification using spectrogram based on multi-channel fusion of dcnn,” *Entropy*, vol. 23, no. 11, Nov. 2021, doi: 10.3390/e23111507.
- [8] R. H. D. Zottesso, Y. M. G. Costa, D. Bertolini, and L. E. S. Oliveira, “Bird species identification using spectrogram and dissimilarity approach,” *Ecol Inform*, vol. 48, pp. 187–197, Nov. 2018, doi: 10.1016/j.ecoinf.2018.08.007.
- [9] J. González, J. Padrón, I. Barbero, L. Custodio, and F. M. Fernando Merchán, “Reconocimiento de canto de aves basado en el análisis de componentes principales del espectrograma,” *Revista de Iniciación Científica*, vol. 5, pp. 124–129, Nov. 2019, doi: 10.33412/rev-ric.v5.0.2398.
- [10] R. Aráuz *et al.*, “CUARTO INFORME NACIONAL DE PANAMÁ ANTE EL CONVENIO SOBRE LA DIVERSIDAD BIOLÓGICA Cuarto Informe Nacional de Panamá ante el Convenio sobre la Diversidad Biológica Contenido”.
- [11] Prensa, “Más de mil especies de aves habitan en el país,” <https://www.miambiente.gob.pa/mas-de-mil-especies-de-aves-habitan-en-el-pais/#:~:text=En%20Panam%C3%A1%20se%20han%20registrado,solo%20existen%20en%20nuestro%20pa%C3%ADs.,> May 09, 2021.
- [12] Santander Universidades, “Python: qué es y por qué deberías aprender a utilizarlo,” <https://www.becas-santander.com/es/blog/python-que-es.html>, Apr. 09, 2021.
- [13] “JavaScript | MDN.” <https://developer.mozilla.org/en-US/docs/Web/JavaScript#reference> (accessed Jun. 09, 2023).
- [14] M. Abadi *et al.*, “TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems,” 2015. [Online]. Available: [www.tensorflow.org](http://www.tensorflow.org).
- [15] nikhil15rana, “¿Cómo instalar Librosa Library en Python?,” <https://www.geeksforgeeks.org/how-to-install-librosa-library-in-python/>, Oct. 06, 2021.
- [16] NumPy community, “NumPy User Guide Written by the NumPy community,” 2022.
- [17] “Welcome to Flask — Flask Documentation (2.3.x).” <https://flask.palletsprojects.com/en/2.3.x/> (accessed Jun. 09, 2023).
- [18] “UTP-Ridda2,” <https://ridda2.utp.ac.pa/>.
- [19] “Google Académico,” <https://scholar.google.es/schhp?hl=es>.
- [20] “CommonKADS.” <https://commonkads.org/> (accessed Jun. 18, 2023).
- [21] L. Roberts, “Understanding the Mel Spectrogram,” <https://medium.com/analytics-vidhya/understanding-the-mel-spectrogram-fca2afa2ce53>, Mar. 05, 2020.
- [22] R. Venkatesan and B. Li, “Convolutional Neural Networks in Visual Computing.”