

## Software de Detección Automática de Fauna y Eventos Sonoros por Correlación Espectral en dos Dimensiones

Sergio Moya\*, José Szychowski

*Facultad de Ingeniería, Universidad Nacional de Misiones (UNaM), Oberá, Misiones, Argentina.*

*GID-IE, FI-UNaM, Oberá, Misiones, Argentina.*

*\*Wildlife Sensor, Leandro N. Alem, Misiones, Argentina*

*sergio.e.moya@gmail.com*

---

### Resumen

En este trabajo se muestran los resultados del desarrollo de un software de detección de biodiversidad en sonidos a pedido del Instituto de Biología Subtropical (IBS), una institución de CONICET y la Universidad Nacional de Misiones. El software cuenta con una interfaz gráfica simple de usar y en la cual, ingresando un patrón o plantilla que contenga el sonido a buscar, realiza la búsqueda de similitudes en archivos de audio, generando diversos resultados configurables. Este software es un desarrollo innovador en el área de la conservación natural y representa una herramienta de gran valor para la generación de estadísticas poblacionales y la detección de especies amenazadas o nuevas para la ciencia.

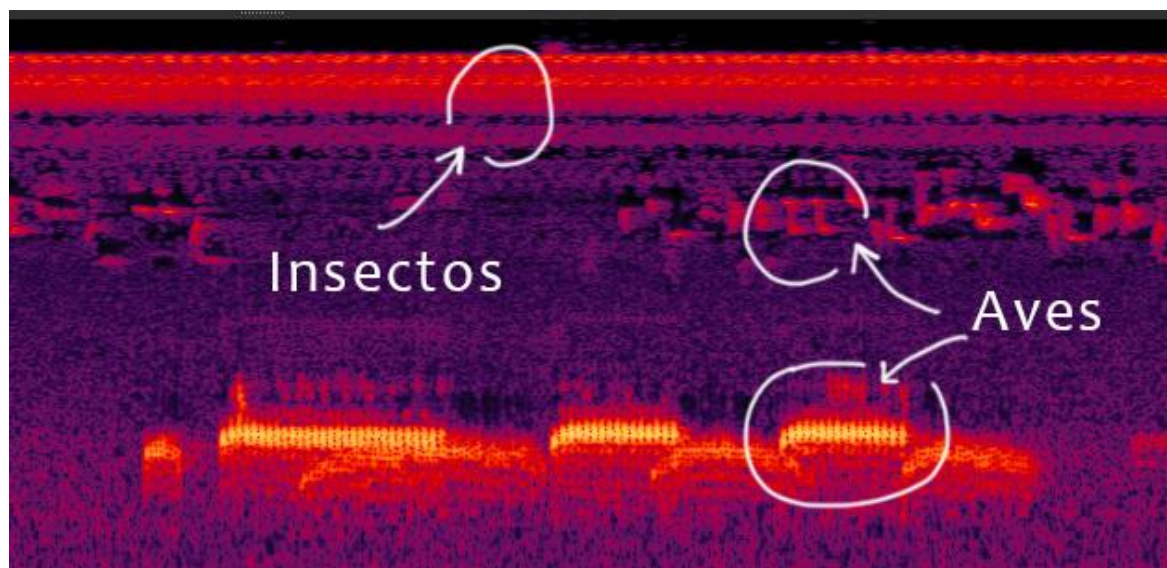
*Palabras Clave – detección de fauna, selva de misiones, monitoreo de biodiversidad.*

### 1. Introducción

La biodiversidad natural y la conservación de ambientes naturales es foco de diversas tareas de investigación en distintas partes del mundo. Una de las áreas más importantes radica en la identificación de especies naturales, las cuales con su presencia permiten inferir la sanidad, o también la degradación, de ciertas áreas naturales. En la provincia de Misiones existe uno de los ambientes naturales más importantes y amenazados del mundo, la Selva Paranaense [1], [2]; en donde habitan un sinnúmero de especies de flora y fauna, algunas de ellas críticamente amenazadas como el Yaguareté (*Phantera onca*) [3], Mono Aullador Rojo (*Alouatta guariba*) [4], Loro Vinoso, Águila Harpía, etc. Debido a esta riqueza natural, Misiones es desde el año 2019 Capital Nacional de la Biodiversidad según la Ley 27494 [5]. En un caso particular, el Instituto de Biología Subtropical (IBS) con sede en Iguazú, organismo dependiente de CONICET y la Universidad Nacional de Misiones (UNaM) posee líneas y proyectos de investigación aplicados a la conservación de dichas especies, y en especial, busca determinar y catalogar la biodiversidad presente en las selvas de Misiones, una tarea de altísima complejidad. Para esto, en los últimos años han avanzado en extensivas campañas de muestreo y relevamiento de datos sobre la biodiversidad de Misiones utilizando grabadores de audio [6] en un gran número de áreas naturales protegidas o amenazadas. Se han adquirido una enorme cantidad de horas de grabación que contienen vocalizaciones de fauna de todo tipo de fauna, ya sea mamíferos, aves, insectos y anfibios, y el objetivo de los investigadores es catalogar dichas especies de manera eficiente. Teniendo en cuenta la imposibilidad de escuchar todas estas horas y listar todo lo oído, se contrató a Wildlife Sensor para desarrollar un software capaz de realizar esta tarea de manera rápida, eficiente y utilizando alguna técnica que permita a cualquier usuario buscar una especie o sonido en particular en esta enorme base de datos de audio. El desafío

\*Sergio Moya

principal consistió en desarrollar una técnica que permita detectar sonidos de todo tipo, muy variantes entre sí y con la mayor eficiencia posible, a la vez de que el tiempo necesario para el procesamiento y búsqueda de los sonidos no sea elevado, es decir, que sea una solución computacionalmente eficiente. La tarea se dividió en dos partes, el desarrollo, simulación y prueba del algoritmo de detección y luego la etapa de desarrollo de una interfaz gráfica que permita al usuario un uso simple del programa, generando reportes claros y concretos que permitan la rápida detección de los eventos sonoros buscados. A modo de ejemplo, la Figura 1 muestra el espectrograma de una porción de 25 segundos de audio en donde se observan distintos cantos de aves e insectos, los cuales deben ser detectados por el programa de manera automática.

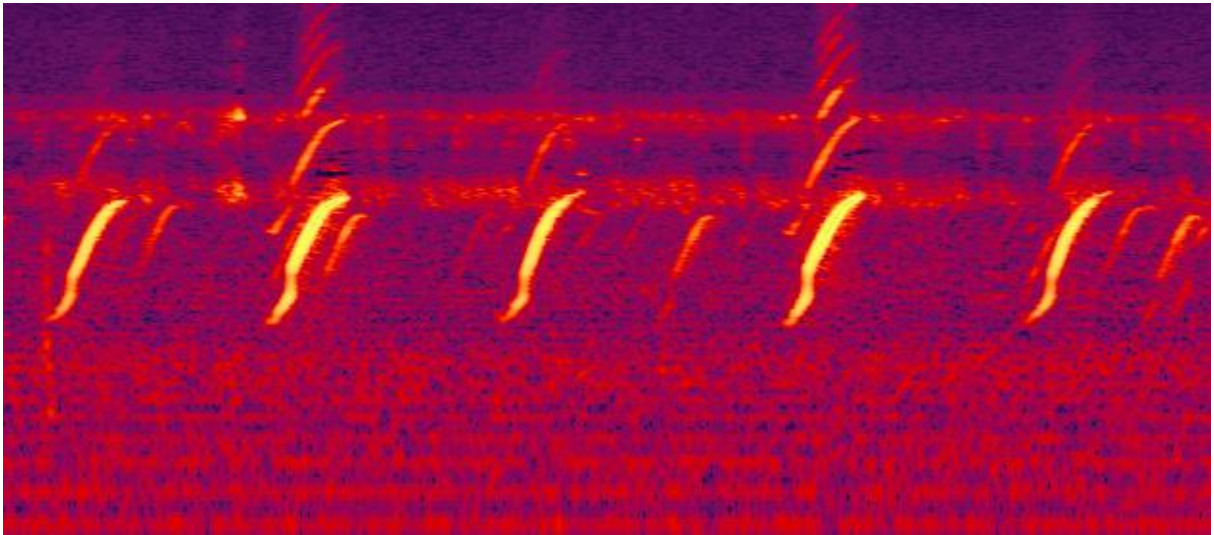


**Fig. 1.** Espectrograma de grabaciones de audio realizadas en la selva. Se observan diversas vocalizaciones de aves y de insectos.

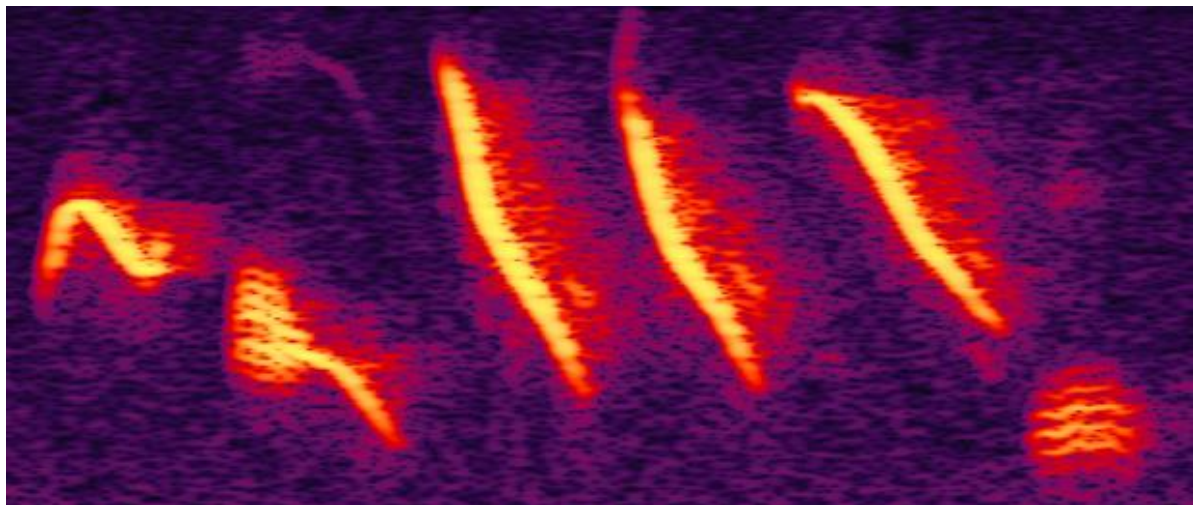
## 2. Desarrollo de algoritmo de detección de sonidos

En base a los requerimientos técnicos y funcionales del software solicitado, se evaluaron diversas alternativas y técnicas posibles que permitan la detección de eventos sonoros que podrían variar enormemente en sus características, tanto frecuenciales como temporales. Luego de diversas pruebas y comparaciones, se optó por desarrollar un algoritmo propio y basado, entre otras cosas, en la correlación de espectrogramas similares al mostrado en la Figura 1, utilizando como plantilla o patrón de comparación, un evento de audio con las características del sonido a buscar ingresado por el usuario. Esta técnica permite entonces encontrar de manera automática eventos sonoros en archivos de audio especificados por el usuario mediante una o diversas plantillas, las cuales son un archivo de audio de corta longitud que contiene la grabación clara del evento sonoro que se desea buscar y detectar. Utilizando esta plantilla, el programa realiza la búsqueda de sonidos similares tanto en tiempo como en frecuencia, generando luego una carpeta con las detecciones logradas en el mismo directorio en donde se buscó el evento. En la Figura 2a se muestra una plantilla perteneciente a la

rana *Leptodactylus fuscus*, relativamente común en la provincia de Misiones, y en la Figura 2b se muestra la plantilla del canto de un ave, en este caso de *Saltator similis*.



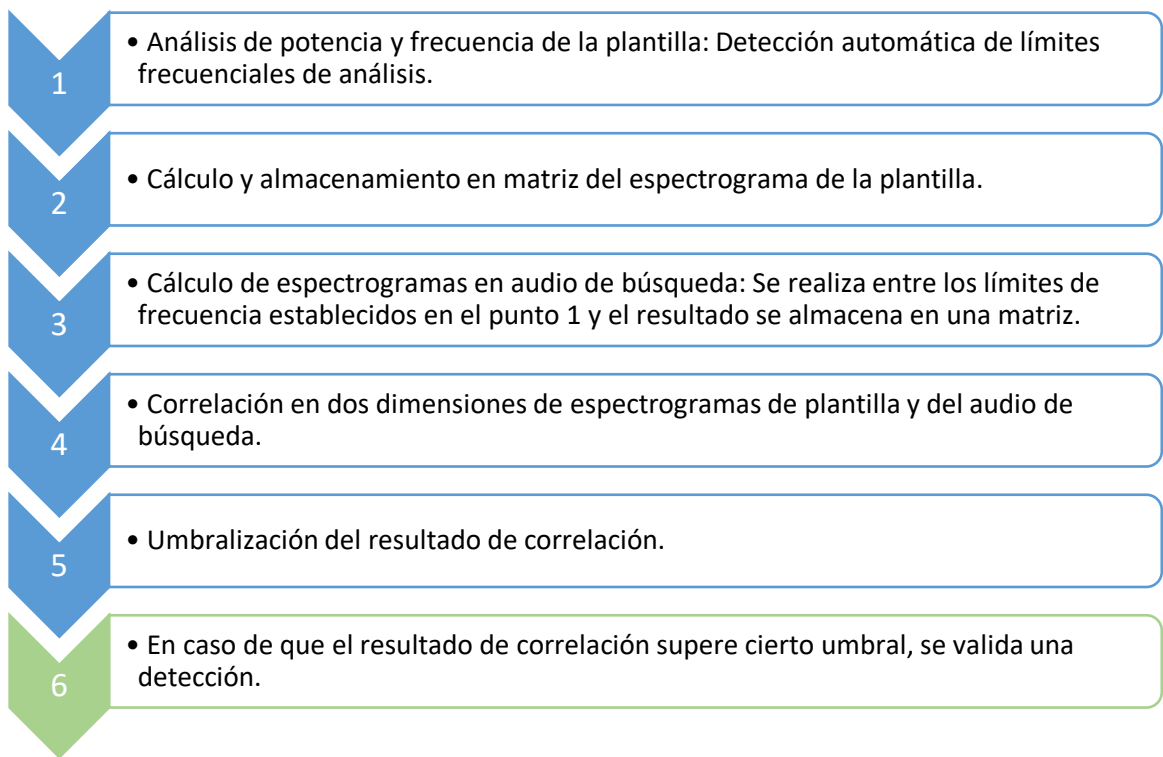
a)



b)

**Fig. 2.** a) Plantilla de la especie de rana *Leptodactylus fuscus*. b) Plantilla de ave *Saltator similis*.

Para realizar la correlación de los espectros de la plantilla y el audio, el programa debe realizar un pre procesamiento de los audios para que el proceso completo sea eficiente en cuestiones computacionales. Un resumen del proceso algorítmico del programa muestra a continuación:



### 3. Desarrollo de interfaz gráfica

Una vez probados los resultados del algoritmo en simulación, se debió desarrollar, en base a los requerimientos del usuario una interfaz gráfica que sea simple y permita trabajar con múltiples carpetas, archivos y modificar parámetros de búsqueda. La interfaz que finalmente fue presentada para el usuario se muestra en la Figura 3:



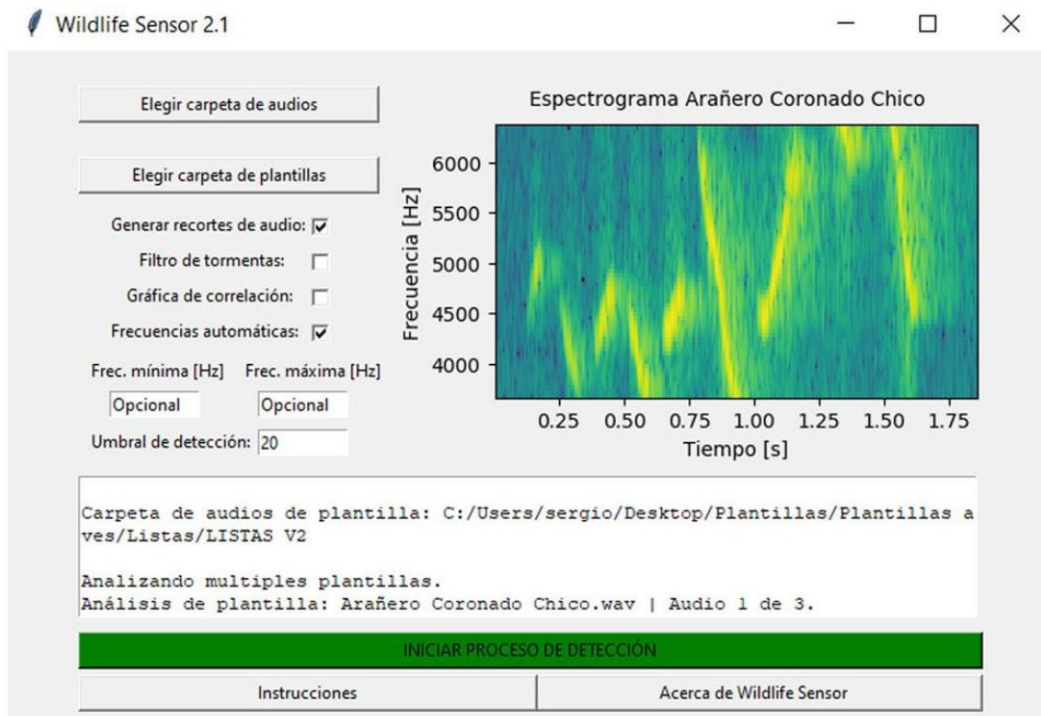
**Fig. 3. Interfaz gráfica desarrollada para el uso práctico del programa.**

La interfaz gráfica consiste en una pantalla única en donde se deben configurar los parámetros para la búsqueda. A su vez posee una ventana de texto informativo en donde se dan avisos sobre el avance del procesamiento o errores que puedan surgir. A continuación se muestran y explican todas las características principales de la interfaz gráfica del programa, como así también el funcionamiento de cada una de sus opciones.

- Elegir carpeta de audios: Haciendo clic en este botón se abre el explorador y debe seleccionarse la carpeta que contiene los archivos de audio en donde se desea buscar el evento sonoro.
- Elegir carpeta de plantillas: En esta opción se debe especificar la carpeta en donde se encuentran las plantillas de audio que se desean buscar. La carpeta puede contener una o más plantillas.
- Generar recortes de audios: Esta opción habilita o no la generación de porciones de audio .wav en las cuales se encuentran las detecciones efectuadas por cada plantilla. Dicho audio es de una duración proporcional a la duración de la plantilla que se utilizó en ese caso, y su nombre es del tipo: Score\_0.28\_AA1\_20191010\_054900\_2\_minuto 7 con 54 segundos.wav. Esta detección, por ejemplo, tiene un Score de 0.28, ocurrió en el audio “AA1\_20191010\_054900”, fue la detección n° 2 de dicho audio y el evento ocurrió en el minuto 7 con 54 segundos.
- Filtro de tormentas: Esta opción habilita o no un filtro que permite eliminar aquellas detecciones que se encuentran completamente saturadas en amplitud y potencia. Esto ocurre principalmente por tormentas, mucho viento o insectos como chicharras, cuya potencia sonora distorsiona completamente el audio y normalmente generan falsas detecciones.
- Gráfica de correlación: Al habilitar esta opción y terminar el procesamiento de una plantilla se muestra un gráfico en pantalla en el cual se grafica la correlación de la plantilla con el ultimo audio procesado, superpuesto por el umbral que se calcula en base al valor ingresado.
- Frecuencias automáticas: El programa calcula automáticamente el rango de frecuencias óptimo que permite analizar cada una de las plantillas, pero en caso de que se observe en pantalla que el rango de frecuencias calculado automáticamente no es correcto, se puede elegir un rango de frecuencias manual, ingresando los mínimos y máximos (en Hertz) en las ventanas inferiores.
- Umbral de detección: Esta es una configuración importante del proceso de búsqueda. Este valor representa un umbral que determina la sensibilidad del programa.

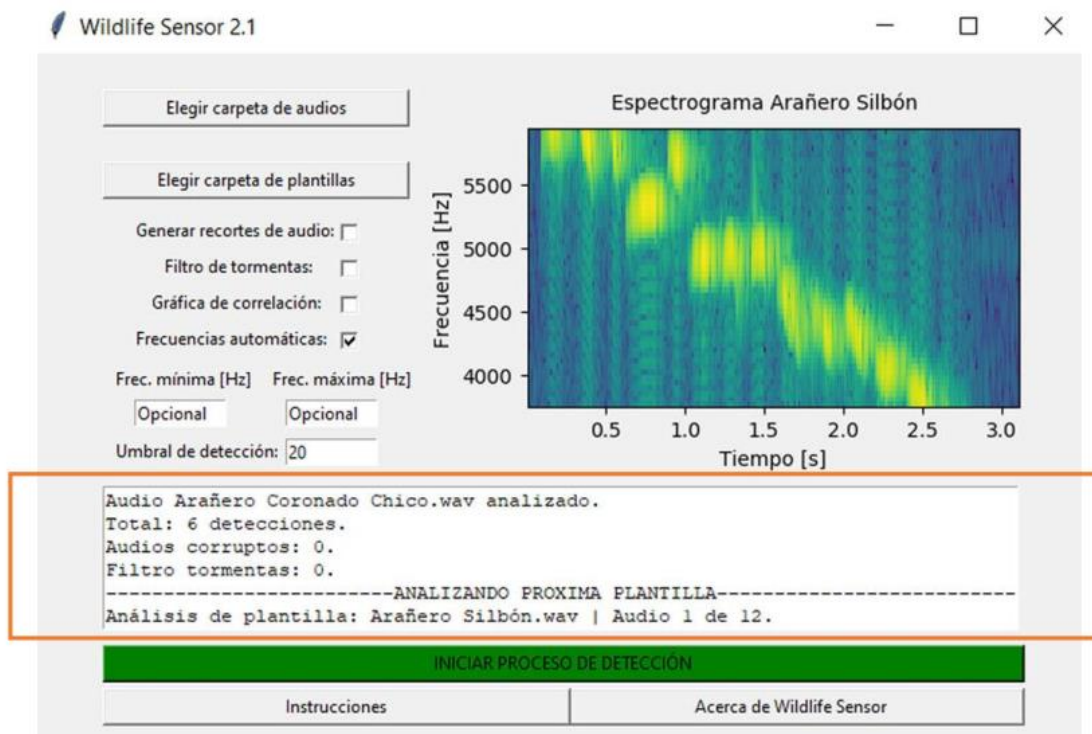
#### **4. Resultados y funcionamiento**

Finalmente, cuando todas estas opciones descriptas anteriormente fueron cargadas y configuradas, se debe hacer clic en el botón verde “Iniciar proceso de detección” (Fig. 3) y en ese momento se comienza analizando la primera plantilla, mostrándose su correspondiente espectrograma y límites de frecuencia calculados automáticamente (en caso de seleccionar dicha opción). Por ejemplo, en la Figura 4, la plantilla analizada corresponde a la especie de ave “Arañero Coronado Chico”, y los límites de frecuencia de análisis son desde los 3500 Hz. (aprox.) hasta los 6500 Hz. (aprox.), y la misma dura aproximadamente 1.8 segundos.



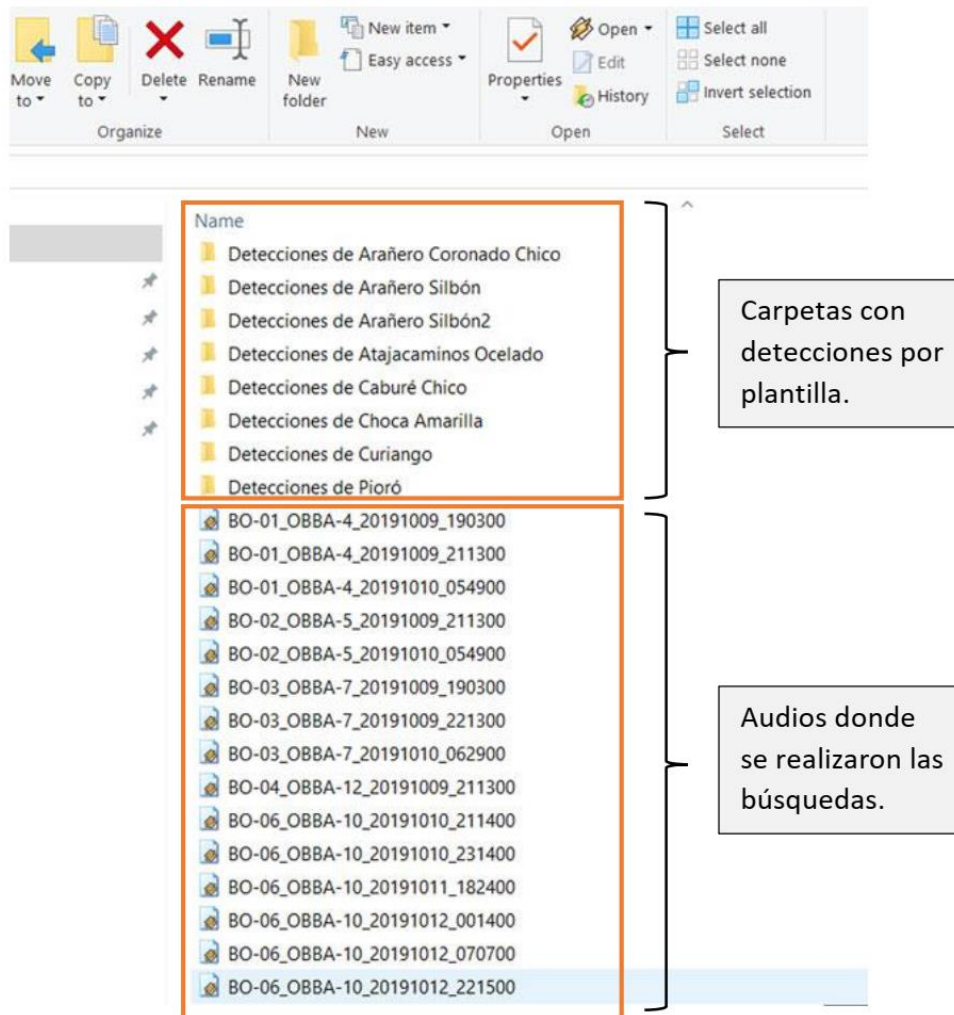
**Fig. 4. Interfaz gráfica en proceso de búsqueda.**

Durante el proceso de detección en el caso de utilizar múltiples plantillas, al finalizar cada una el software muestra en pantalla un resumen del procesamiento de cada plantilla, reportando el número de detecciones, audios corruptos o filtrados. Esto se puede ver en la Figura 5.



**Fig. 5. Reporte de resultados en pantalla entre análisis de plantillas.**

Los archivos detectados (en caso de que haya alguno) se graban en una carpeta creada con tal fin y denominada “Detecciones de *NombreDePlantilla*”. Estas carpetas se encontrarán en la misma carpeta donde estaban los audios en donde se buscó el evento e indicada en la ventana de notificaciones del programa. La Figura 6 muestra el resultado de búsqueda de diversas plantillas:



**Fig. 6. Carpeta final donde se guardan las detecciones**

Dentro de cada carpeta de detección se encontrarán los archivos de audio con dichas detecciones (en caso de que se haya elegido generarlos) y un archivo de texto TXT con los detalles de configuración de la búsqueda y una lista de todas las detecciones, ordenadas de mayor a menor “score”. La Figura 7 muestra un ejemplo de este archivo de texto. En este informe se documenta cada configuración seleccionada y se especifica el valor exacto de rangos de frecuencia analizados, que para el caso del ejemplo, correspondiente a la plantilla “Atajacaminos Ocelado”, fue entre 1291 Hertz y los 1937 Hertz, hubo 206 detecciones y el tiempo que se tardó en procesar dicha plantilla fue de 54.09 segundos (dato al final del archivo .txt).

```

Informe - Notepad
File Edit Format View Help
detecciones para plantilla: "Atajacaminos Ocelado.wav"

Fecha y hora del analisis: 04/08/20 16:58:35
Carpeta de analisis: C:/Users/sergio/Desktop/audio pruebas

Total de detecciones: 206
Valor de umbral: 20.0

Filtro de tormentas: No
Generar recortes de audio: No

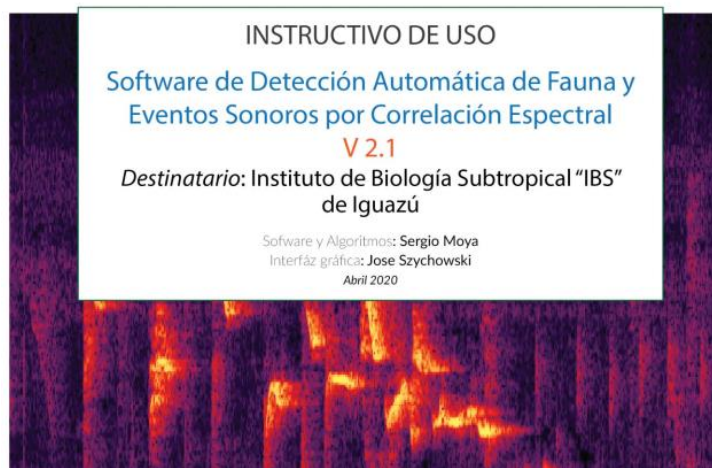
Frecuencia minima de analisis: 1291 Hertz
Frecuencia maxima de analisis: 1937 Hertz

Nombre Audio                               Tiempo Score
1 BO-01_OBBA-4_20191009_211300.wav         minuto 5 con 33 segundos 0.59
2 BO-03_OBBA-7_20191009_190300 .wav       minuto 9 con 46 segundos 0.59
3 BO-06_OBBA-10_20191010_211400 .wav     minuto 9 con 11 segundos 0.59
4 BO-01_OBBA-4_20191009_211300.wav         minuto 4 con 28 segundos 0.52
5 BO-01_OBBA-4_20191009_211300.wav         minuto 6 con 11 segundos 0.52
6 BO-06_OBBA-10_20191010_211400 .wav     minuto 7 con 49 segundos 0.52
7 BO-03_OBBA-7_20191009_190300 .wav       minuto 9 con 3 segundos 0.52
8 BO-06_OBBA-10_20191010_211400 .wav     minuto 9 con 18 segundos 0.52
9 BO-06_OBBA-10_20191010_211400 .wav     minuto 9 con 15 segundos 0.50
10 BO-01_OBBA-4_20191009_211300.wav        minuto 5 con 55 segundos 0.50
11 BO-01_OBBA-4_20191009_211300.wav        minuto 6 con 25 segundos 0.50
12 BO-06_OBBA-10_20191010_211400 .wav     minuto 9 con 23 segundos 0.50

```

**Fig. 7. Archivo TXT de informe de resultados de detección.**

Finalmente se escribió un manual de uso (Fig. 8) en donde se especifican al detalle todas las consideraciones técnicas que deben reunirse para el correcto funcionamiento del programa, como así también como hacerlo funcionar de manera más eficiente; este documento se entregó a los investigadores del IBS.



**Fig. 8. Carátula de manual de uso entregado a los investigadores del IBS.**



## 5. Conclusiones

Este trabajo presentó un desarrollo de un software a pedido de una institución de investigación de la provincia de Misiones. El objetivo del software es detectar automáticamente y de manera eficiente sonidos específicos (y muy variados) en grabaciones de audio que pueden tener miles de horas de duración total con el objetivo de listar y detectar la biodiversidad presente en la provincia de Misiones. Para esto, el proceso de desarrollo se nutrió de etapas de investigación de algoritmos, simulaciones y el vínculo con colegas especialistas en desarrollo de software específico para la interfaz gráfica, logrando de esta manera un producto fácil de usar y que reunió todas las condiciones requeridas por el cliente.

## Agradecimientos

Se agradece al Departamento de Electrónica y la Facultad de Ingeniería de la Universidad Nacional de Misiones (UNaM). Sergio Moya agradece a José Szychowski por sus valiosos aportes y dedicación en las etapas de desarrollo de la interfaz gráfica y entre ambos agradecen a Diego Varela, Investigador de CONICET y miembro del Instituto de Biología Subtropical (IBS) por la confianza puesta en nuestro trabajo y el constante incentivo en el desarrollo de nuevos desafíos.

## Referencias

- [1] Gobierno de la Provincia de Misiones, Ministerio de Ecología y RNR, “Monumentos Naturales,” 2012. [Online]. Available: <http://www.ecologia.misiones.gov.ar/ecoweb/index.php/andescgen/clasificacion/monumentos-naturales>.
- [2] Gobierno de la Provincia de Misiones; Ministerio de Ecología., “Sistemas de Áreas Naturales Protegidas (SANP),” 2017. [Online]. Available: <http://www.ecologia.misiones.gov.ar/ecoweb/index.php/andescgen/sistema-de-areas-naturales-protegidas>.
- [3] Agustín Paviolo, Anne Sophie Bertrand, Angel Caradonna, Apolonio Rodrigues, and Carlos De Angelo, *Plan de Acción para la Conservación de la Población de Yaguareté (Panthera onca) del corredor Verde de Misiones*. Puerto Iguazú: Comisión de conservación del yaguareté de la Selva Paranaense, 2011.
- [4] “The IUCN Red List of Threatened Species,” *IUCN Red List of Threatened Species*. [Online]. Available: <https://www.iucnredlist.org/en>. [Accessed: 24-Jul-2019].
- [5] Boletín Oficial: 3 de enero de 2019, “Ley 27494 – Declaración a la provincia de Misiones Capital Nacional de la Biodiversidad,” de Enero-2019. [Online]. Available: <http://argentinambiental.com/legislacion/nacional/ley-27494-declaracion-la-provincia-misiones-capital-nacional-la-biodiversidad/>.
- [6] Wildlife Acoustics, “Song Meter SM4,” *The smallest and lightest dual-channel weatherproof acoustic recorder available*, 2017. [Online]. Available: <https://www.wildlifeacoustics.com/products/song-meter-sm4>.