

**ANÁLISIS DE PAISAJES ACÚSTICOS**

**LUIS CARLOS RESTREPO FORERO**

**JEFERSON DAVID VARGAS TOCA**

**Trabajo de grado presentado para optar al título de Ingeniero Electrónico.**

**DIRECTOR:**

**PEDRO RAÚL VIZCAYA GUARÍN, Ph.D.**

**CO-DIRECTOR:**

**OSCAR ALBERTO LAVERDE RODRIGUEZ, Ph.D.**

**PONTIFICIA UNIVERSIDAD JAVERIANA**

**FACULTAD DE INGENIERÍA**

**DEPARTAMENTO DE ELECTRÓNICA**

**CARRERA DE INGENIERÍA ELECTRÓNICA**

**2017**



## **AGRADECIMIENTOS.**

Luis Carlos:

A mi mamá, mi papá, mi hermano y mi familia, por haber sido el apoyo cuando lo necesité.

David:

A mi mama, mis hermanos y mi hijo por haber sido el motor que impulsa mi vida.

# TABLA DE CONTENIDO

INTRODUCCIÓN .....	5
CAPITULO 2: MARCO TEÓRICO .....	7
2.1 TRANSFORMADA RÁPIDA DE FOURIER.....	7
2.2 ESPECTRO DE FRECUENCIAS DE UNA SEÑAL .....	7
2.2.1 Análisis espectral.....	7
2.2.2 Espectrograma .....	7
2.3 ENTROPÍA .....	8
2.4 CEPSTRUM .....	8
2.5 RECONOCIMIENTO DE IRIS .....	8
2.6 HUELLA DIGITAL ACÚSTICA .....	8
2.7 KALEIDOSCOPE PRO .....	9
2.8 AUDACITY .....	9
2.9 R.....	9
2.10 MATLAB.....	9
2.10.1 Máquinas de vectores de soporte .....	9
2.10.2 Red Neuronal Artificial .....	10
2.11 DISTANCIA EUCLIDIANA .....	11
2.12 INTERPOLACIÓN DE DATOS.....	11
CAPITULO 3: OBJETIVO DEL PROYECTO.....	12
3.1 OBJETIVO GENERAL.....	12
3.2 OBJETIVOS ESPECÍFICOS .....	12
CAPITULO 4: DESARROLLO.....	13
4.1 BASE DE DATOS .....	13
4.2 PRE-PROCESAMIENTO .....	14
4.3 EXTRACCIÓN DE CARACTERÍSTICAS .....	15
4.3.1 Trabajo en R.....	15
4.3.2 Remodelación de vectores .....	15
4.3.3 Cepstrum .....	16
4.4 DEFINICIÓN DE HUELLA DIGITAL ACÚSTICA.....	17
4.4.1 Interpolación de datos.....	17
4.4.2 Distancia Euclidiana.....	18
4.4.3 Comparación de Huellas digitales acústicas diarias entre clases.....	19
4.5 CLASIFICACIÓN.....	21
4.5.1 Máquinas de vectores de soporte .....	21
4.5.2 Red de reconocimiento de patrones (patternet).....	22
CAPITULO 5: PROTOCOLO DE PRUEBAS .....	25
5.1 DISTANCIA EUCLIDIANA.....	25
5.2 MÁQUINAS DE VECTORES DE SOPORTE.....	25
5.3 RED DE RECONOCIMIENTO DE PATRONES (PATTERNET) .....	25
CAPITULO 6: ANÁLISIS DE RESULTADOS .....	26
6.1 DISTANCIA EUCLIDIANA.....	26
6.2 MÁQUINAS DE VECTORES DE SOPORTE.....	26

6.3 RED DE RECONOCIMIENTO DE PATRONES (PATTERNET) .....	26
CAPITULO 7: CONCLUSIONES Y RECOMENDACIONES .....	27
ANEXOS .....	28
ANEXO 1 .....	28
ANEXO 2 .....	28
ANEXO 3 .....	29
ANEXO 4 .....	30
ANEXO 5 .....	30
ANEXO 6 .....	31
ANEXO 7 .....	32
ANEXO 8 .....	32
ANEXO 9 .....	32
ANEXO 10 .....	32
ANEXO 11 .....	32
ANEXO 12 .....	32
ANEXO 13 .....	32
ANEXO 14 .....	32
BIBLIOGRAFÍA .....	33

# INTRODUCCIÓN

El hombre, al convertirse en una especie sedentaria, ha buscado lugares dónde establecerse (Romero, 2009). También, a medida que crecieron las poblaciones asentadas en zonas urbanas, se presentó un crecimiento de esas urbes sobre las zonas naturales. Esta invasión que ha llevado a cabo el hombre a lo largo del tiempo ha significado grandes cambios en los comportamientos de los diferentes ecosistemas (plantas, animales, climas, etc.).

En ecología se buscan herramientas que le permitan al investigador identificar comportamientos en los diferentes ecosistemas para hallar las diferencias o cambios en ellos y permitirles intervenirlos a tiempo. La mayoría de las herramientas utilizadas actualmente para estudiar los ecosistemas (especies, episodios aislados, comportamientos de las diferentes especies), aparte de significar grandes esfuerzos (desplazamientos, intervenciones a los ecosistemas), también significan una inversión mucho mayor (Sueur, Pavoine, Hamerlynck, & Duvail, 2008).

Los sonidos son una propiedad fundamental de los paisajes (Schafer, 1977) y entregan características importantes para el análisis de los mismos (Pijanowski, Villanueva-Rivera, et al., 2011). Es por eso que el estudio de los sonidos de los pasajes ha venido convirtiéndose en una herramienta de gran uso para los ecólogos, ya que, por no invadir los paisajes y conservar su naturaleza (Sueur, Farina, Gasc, Pieretti, & Pavoine, 2014), se convierte en una práctica ideal ya que, se obtiene información mucho más real de lo que está ocurriendo en el paisaje (Pijanowski, Villanueva-Rivera, et al., 2011). Es importante tener en cuenta la fidelidad de la información ya que, a partir de ahí, se puede dar un diagnóstico más acertado y se puede proponer una solución más efectiva (Caycedo-Rosales, Ruiz-Muñoz, & Orozco-Alzate, 2013).

Para analizar los ecosistemas, los ecólogos acostumbran realizar avistamientos y mantenerse por largo tiempo en salidas de campo para obtener datos del lugar que desean estudiar. En esas salidas de campo se dejan grabadoras para obtener datos adicionales y monitorear lo que ocurre. Normalmente, se realizan estas salidas de campo para estudiar una especie en particular (casi siempre, una especie difícil de avistar), pero, para contextualizar este estudio, necesitan caracterizar el paisaje, escuchando todos los audios y suponiendo diferentes características y comportamientos, como se describió anteriormente (Pijanowski, Villanueva-Rivera, et al., 2011). Este proceso es conocido como *Ecología de Paisajes Acústicos (Soundscape Ecology)* (Pijanowski, Villanueva-Rivera, et al., 2011) y consiste en el análisis de las características acústicas de un ecosistema. El problema para los ecólogos consiste en que, al mantener esas grabadoras por mucho tiempo, revisar la cantidad de información obtenida se convierte en una tarea bastante dispendiosa.

Actualmente, aunque se está empezando a implementar una manera de automatizar este proceso, no es una práctica común de los ecólogos (Pijanowski, Farina, Gage, Dumyahn, & Krause, 2011). También, aunque algunos ecólogos están automatizando el proceso de análisis de paisajes acústicos, lo están haciendo enfocados en obtener ciertas características de algunas especies que habitan en el lugar que están estudiando, en analizar el impacto de la mano del hombre en ese lugar y en obtener índices de las especies grabadas (Kasten, Gage, Fox, & Joo, 2012)

A través de los años, la mano del hombre ha intervenido los ecosistemas, modificando sus condiciones naturales y, en la mayoría de los casos, perjudicando a las diferentes especies que habitan en estos. Los ecólogos buscan la manera de mejorar los ecosistemas y para esto deben analizar su comportamiento. Por medio de los sonidos de los diferentes paisajes se pueden analizar varias características, clasificar los

paisajes e identificar especies; pero este proceso requiere analizar cada segundo de todos los sonidos recopilados y de esta forma identificar cada una de las cosas que se encuentren en los paisajes que se están analizando, permitiendo el espacio para que ocurran errores o se omitan detalles.

Por medio del análisis de señales de los paisajes acústicos se pretende simplificar la tarea de los ecólogos, ya que se podrán analizar todos los registros de audio de los paisajes sin tener que escucharlos en su totalidad. Así, las tareas se realizarían en un proceso mucho más corto y se automatizaría el proceso de preselección según lo que desee obtener el usuario (clasificación del paisaje, especies que allí habitan, etc.). El objetivo general de este trabajo de grado es definir, analizar y evaluar características de paisajes acústicos de bosques del Magdalena Medio Colombiano que permitan su clasificación automática.

# CAPITULO 2: MARCO TEÓRICO

En cuanto al método a utilizar, se parte de la idea que la información contenida en los audios describe un proceso estocástico estacionario en sentido amplio. Para que este proceso sea estacionario en sentido amplio solo se requiere que los dos primeros momentos sean iguales, es decir, que no varíen con respecto al tiempo. Los procesos estacionarios en sentido amplio tienen restricciones en la media y en la función de correlación como lo muestran las ecuaciones (1) y (2) respectivamente.

$$E\{x(t)\} = m_x(t) = m_x(t + \tau) \quad (1)$$

$$E\{x(t_1)x(t_2)\} = R_x(t_1, t_2) = R_x(t_1 + \tau, t_2 + \tau) = R_x(t_1 - t_1, 0) \quad (2)$$

Lo que quiere decir que la media debe ser constante y que la función de correlación depende de la diferencia de los dos momentos  $t_1$  y  $t_2$ . Esta suposición se hace con base al teorema de Wiener-Khinchin (Barkat, 2005). Es importante tener en cuenta esta definición de procesos estacionarios en sentido amplio ya que, es la base de este trabajo de grado. Para archivos de audio que no describan un proceso estacionario, se deben convertir en uno de forma tal que todos los archivos tengan las mismas propiedades. Para que todos los momentos de los archivos de audio describan procesos estacionarios se debe dividir el archivo de audio original en pequeñas partes que describan un mismo momento.

## 2.1 Transformada rápida de Fourier

Es un algoritmo que permite hacer el cálculo de la transformada de Fourier a una señal, para pasarla al dominio de la frecuencia si originalmente estaba en el dominio del tiempo. El resultado, al obtener esa representación en el dominio de la frecuencia, permite obtener la magnitud y la fase de una señal en el dominio de la frecuencia. Sirve para conocer en qué frecuencias de la señal hay una mayor intensidad. La transformada rápida de Fourier es la aplicación de la transformada de Fourier a señales discretas (Bergland, 1969). Según estándares, para la obtención de una huella digital acústica, uno de los primeros pasos es el análisis del espectro de frecuencias que se puede obtener al aplicar la Transformada rápida de Fourier a un audio. Para este trabajo de grado, se usó la transformada rápida de Fourier para calcular el Cepstrum de las señales a trabajar.

## 2.2 Espectro de frecuencias de una señal

El espectro en frecuencia es la distribución de las amplitudes de un fenómeno ondulatorio que sea superposición de ondas de varias frecuencias. Debido a que las ondas están a diferentes frecuencias, el espectro de frecuencias muestra la intensidad frente a la frecuencia a la que ocurre. Para caracterizar mejor un archivo de audio es necesario hacer un análisis espectral (Fabbri & Abagnali, 2012). Para hallar el espectro en frecuencias de una señal se le aplica la transformada de Fourier.

### 2.2.1 Análisis espectral

El análisis espectral cuantifica las intensidades de los sonidos de un archivo de audio. Este proceso se lleva a cabo al descomponer el archivo de audio por partes, ubicando los sonidos en el tiempo y mostrando la frecuencia y amplitud que tiene cada sonido. El análisis espectral es representado gráficamente en un espectrograma (Fabbri & Abagnali, 2012).

### 2.2.2 Espectrograma

El espectrograma caracteriza los sonidos de una señal de audio mostrando la amplitud y la frecuencia que tiene cada sonido ubicándolo en el tiempo en el que ocurre (Fabbri & Abagnali, 2012). Un ejemplo de un espectrograma se puede encontrar en la Figura 2.

### 2.3 Entropía

La entropía entrega la cantidad de información que tiene una fuente. Para el caso de los archivos de audio, la entropía entregará la información que contiene el audio sobre la clase de bosque. La entropía, como la calcula R, está definida como lo muestra la ecuación (3).

$$H = -\frac{\sum_{i=1}^N y_i \log_2(y_i)}{\log_2(N)} \quad (3)$$

### 2.4 Cepstrum

El Cepstrum es el cambio en el espectro de una señal. El Cepstrum de una señal es el resultado de calcular la transformada inversa de Fourier al espectro de una señal en escala logarítmica (Oppenheim & Schaffer, 2004). Suponiendo una señal  $x[t]$ , el Cepstrum está definido como lo muestra la ecuación (4).

$$F^{-1}\{\log(|F\{x[t]\}|)\} \quad (4)$$

Donde  $F$  es la transformada de Fourier y  $F^{-1}$  es la transformada inversa de Fourier (Oppenheim & Schaffer, 2004).

El Cepstrum entregará información de variaciones del espectro de una señal, conteniendo la información de las variaciones lentas en los coeficientes iniciales. Suponiendo que las dos clases de paisajes son diferentes (en cuanto a las características que muestran los audios), los coeficientes iniciales del Cepstrum nos ayudaran a la clasificación ya que los cambios en los paisajes son lentos (durante todo el día).

Es importante el cálculo del Cepstrum ya que este ayudará a la extracción de características de los audios en un determinado momento y así poder conformar la huella digital acústica de cada clase.

### 2.5 Reconocimiento de iris

El reconocimiento de iris es un de las técnicas con mayor importancia para el reconocimiento de patrones por el método utilizado. El iris tiene una gran dificultad para hacerle un reconocimiento por su forma circular. La técnica usada para hacer el reconocimiento, luego de estandarizar los tamaños del iris, conformó una representación rectangular del iris, una forma matricial que caracterice el mismo. La técnica de reconocimiento de iris subdivide esa representación rectangular en pequeñas porciones. A cada una de esas porciones de imagen se le saca un valor promedio que representa el valor de cada pequeña porción de la imagen original (Huang, Luo, & Chen, 2002). Al final, al hacer estas divisiones y obtener el promedio de las mismas, se reduce la dimensión de la información original.

Con base a esta técnica se redujo la información original de los archivos de audio al subdividir el vector original en pequeños vectores, agruparlos en una matriz, calcularle el Cepstrum y obtener la media, por columna, de los signos. La extracción de estos datos ayudará a la caracterización de cada archivo de audio para la posterior conformación de la huella digital acústica de cada clase.

### 2.6 Huella digital acústica

Es la información única de un lugar basado en la información de los sonidos del lugar. La huella digital acústica es el patrón de audio que tiene un lugar cuando se presentan comportamientos característicos del mismo. La aplicación más importante de la huella digital acústica es para el reconocimiento de patrones (Koutroumbas, S. TheoDoridis, 2003). La huella digital acústica entrega información de la clase a la cual pertenece, sin embargo, esta información no es necesaria para una clasificación. Si se obtiene la huella digital acústica de una clase específica (lugar, ambiente, etc.) se podría obtener el patrón de audio de la misma, por lo tanto, si se tienen fragmentos de audio de cualquier otra clase, al comparar ese



fragmento con la huella de la clase inicial, se podrá determinar si pertenece o no a la clase de la que se obtuvo la huella.

Los siguientes programas fueron usados para llevar a cabo este trabajo de grado.

### **2.7 Kaleidoscope Pro**

Software que, al igual que Audacity, le permite al usuario trabajar con archivos de audio. Se centra en el trabajo de los ecólogos enfocados en audio. Por la necesidad de caracterizar los sonidos emitidos por los animales, también le permite al usuario sacar el espectrograma de los archivos de audio. Kaleidoscope Pro, pensando en los usuarios que hacen análisis de los archivos de audio, provee un convertidor de archivos, ya que la mayoría de grabadoras comprimen los audios en formato WAC para ocupar menos memoria. Es utilizado en este trabajo de grado por la necesidad de convertir los archivos de audio de formato WAC a formato WAV, ya que los programas R y MATLAB trabajan con este formato de archivo.

### **2.8 Audacity**

Es un software que permite trabajar con archivos de audio. Le permite al usuario disminuir la velocidad de reproducción, ver la forma de onda de la señal de audio, obtener el espectrograma de archivos de audio entre otras funciones. Es utilizado en este trabajo de grado para hacer un análisis inicial de los archivos de audio a partir de sus datos tales como su espectrograma.

### **2.9 R**

Es un programa con un lenguaje de programación propio. Está enfocado al análisis estadístico, aunque, por ser un software libre, es ampliamente utilizado por biólogos, economistas y profesionales de otras áreas, dado que tiene la capacidad de cargar paquetes o bibliotecas correspondientes al trabajo de diferentes áreas y objetivos (por ejemplo, el paquete *Soundecology* para ecólogos, o el paquete *Value at Risk* para economistas) (Leemis, 2016). Es utilizado en este trabajo de grado para obtener índices, que permitan la caracterización de los paisajes, a partir de los archivos de audio (Villanueva-Rivera & Pijanowski, 2016).

### **2.10 MATLAB**

Software matemático con lenguaje de programación propio. MATLAB es un programa que permite trabajar en diferentes tipos de aplicaciones, brindando al usuario diferentes herramientas, funciones y recursos para llevar a cabo gráficas, simulaciones, aplicaciones, etc., o para simplemente solucionar problemas comunes. Por otro lado, MATLAB, tiene adiciones (llamadas toolboxes o cajas de herramientas), que le permiten al usuario trabajar en diferentes áreas, como, por ejemplo, la adición de procesamiento de señales, la adición de las redes neuronales artificiales, etc. (Amos Gilat, 2012).

#### **2.10.1 Máquinas de vectores de soporte**

Las Máquinas de vectores de soporte, o SVM por sus siglas en inglés, son un conjunto de métodos de aprendizaje supervisado. Es un modelo que permite etiquetar en dos clases un conjunto de muestras luego de un entrenamiento. El entrenamiento se hace por medio de un conjunto de muestras iniciales de las dos clases debidamente etiquetadas. El objetivo será entonces que, con muestras singulares de alguna de las dos clases, el SVM sea capaz de clasificar entre una clase y otra.

En cuanto al entrenamiento, SVM hace la separación de clases por medio de un hiperplano que corresponde al vector de separación de los puntos, de diferente clase, que se encuentren más cercanos en todo el Plano, dividiendo entre las dos clases al pasar por la mitad de la distancia de separación de las

muestras menos separadas de clases diferentes. Por su parte, MATLAB, al aplicar SVM sobre dos clases de datos, entrega un margen que corresponde a la suma de la distancia de separación entre el hiperplano y los puntos más cercanos de clases diferentes, como lo muestra la Figura 2.

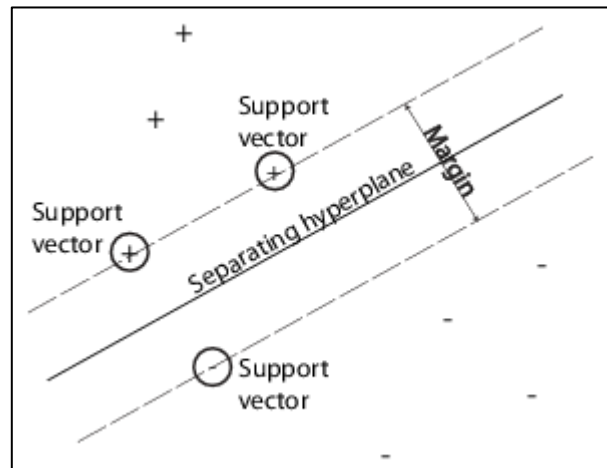


Figura 2. Procedimiento que hace MATLAB al aplicar SVM a las muestras de dos clases e datos. (MATLAB, 2016b)

Este método es ideal para la separación de clases binarias de una base de datos.

### 2.10.2 Red Neuronal Artificial

Es un sistema computacional basado en el conjunto de unidades neuronales simples, emulando, de forma aproximadamente analógica el comportamiento de la conexión de las neuronas (Haykin, 1994). Por lo mismo, son una técnica de aprendizaje y procesamiento automático. Las neuronas, que componen la red neuronal, almacenan conocimiento experimental para dejarlo a disposición del usuario posteriormente (Haykin, 1994).

Teniendo en cuenta que las neuronas aprenden de ellas mismas y de la información que les entregan las otras neuronas, sería correcto decir que la red neuronal tendrá un mejor aprendizaje a medida que se tengan más conexiones entre las neuronas, es decir, mayor cantidad de neuronas. A medida que se hagan más pruebas con la información de entrada, la red neuronal irá cambiando sus características y ajustará sus respuestas para una mejor clasificación, por lo tanto, entre más veces se haga el entrenamiento (iteraciones) mejor será la clasificación de la entrada de la red. Para realizar la clasificación, la red neuronal realiza un entrenamiento, dependiendo de las iteraciones que ingrese el usuario, para, finalmente, realizar un proceso de comprobación que la información ingresada haya sido correctamente clasificada (pruebas). A medida que la cantidad de errores en el entrenamiento y en el proceso de pruebas sea más cercano se tendrá una clasificación más acertada de los datos ingresados. Cuando los errores de entrenamiento y de prueba son iguales se tendrá que la red neuronal está sobreentrenada, es decir, que la cantidad de neuronas en las capas es mayor a la necesaria.

La aplicación de esta herramienta en MATLAB entrega gran cantidad de información para tener un criterio claro sobre los resultados obtenidos y para hacer un análisis pertinente sobre los resultados. MATLAB entrega información de antes y después del entrenamiento e información luego de haber realizado las pruebas de la clasificación hecha en la Red Neuronal.

#### 2.10.2.1 Red de reconocimiento de patrones (patternet)

Son un tipo de red neuronal de tipo Feedforward que pueden ser entrenadas para clasificar las entradas de acuerdo a las clases que se desean a la salida. Se le entrega la información de las etiquetas de cada entrada a la red para que encuentre un patrón y clasifique entre clases. Esta red neuronal está

compuesta por una capa oculta y una capa de salida (MATLAB, 2016a). Es una red ideal para clasificación de datos y el reconocimiento de patrones.

Las siguientes definiciones o procedimientos son necesarios para completar este trabajo de grado.

### 2.11 Distancia euclidiana

Es la distancia entre dos puntos, calculada a partir del teorema de Pitágoras. La representación de la distancia euclidiana se puede ver en la Figura 1.

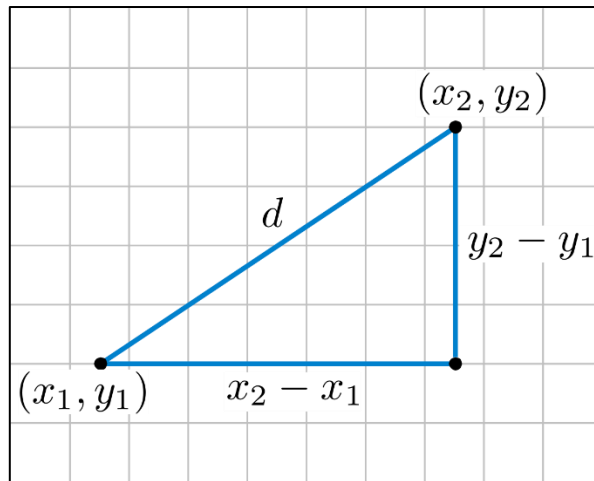


Figura 1. Representación gráfica de la distancia euclidiana.

donde  $d$  es la distancia y se calcula por medio de la ecuación (5).

$$d(P_1, P_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (5)$$

En MATLAB, la distancia euclidiana entre matrices opera por filas, obteniendo la raíz cuadrada de la suma de los cuadrados de la resta de elemento por elemento de la fila, depositando este valor en la posición  $i, j$ , correspondientes a la fila  $i$  de la matriz 1 y la fila  $j$  de la matriz 2.

### 2.12 Interpolación de datos

Es un procedimiento matemático que se realiza para obtener de nuevos puntos de datos partiendo de la idea que se conoce el comportamiento del resto de datos. La aplicación de la interpolación de datos es la obtención de elementos que pertenezcan y vayan de acuerdo al comportamiento del resto de datos. Aunque es una herramienta rápida y sencilla, no siempre es muy precisa. Otra aplicación de la interpolación puede ser completar los datos de una base de datos incompleta.

## CAPITULO 3: OBJETIVO DEL PROYECTO

El objetivo general de este trabajo de grado es definir, analizar y evaluar características de paisajes acústicos de bosques del Magdalena Medio Colombiano que permitan su clasificación automática. En este trabajo de grado se logró obtener un método de extracción de características. Sin embargo, estas características extraídas no representaron un patrón para realizar una clasificación.

### **3.1 Objetivo general**

Definir, analizar y evaluar características de paisajes acústicos de bosques del Magdalena Medio Colombiano que permitan su clasificación automática.

### **3.2 Objetivos específicos**

- Procesar el audio para extraer las características de cada clase de paisaje y analizar su clase.
- Obtener la huella de las dos clases de paisajes.
- Clasificar los audios con base a las clases escogidas y las características seleccionadas, y evaluar su desempeño.
- Realizar pruebas sobre dos tipos de bosques diferentes del Magdalena Medio Colombiano que corresponden a un bosque inundable y un bosque no inundable. Esto se realiza para obtener la huella de cada clase de bosque y así identificarlos fácilmente.

# CAPITULO 4: DESARROLLO

En general, el desarrollo que se pensaba llevar a cabo consistía al procedimiento descrito por la Figura 2, donde, luego de hacer un preprocesamiento de la base de datos, se extraerían las características y se conformaría la huella digital, se clasificarían los datos obtenidos por cada clase y, finalmente, se realizarían pruebas y la validación de las mismas pruebas.

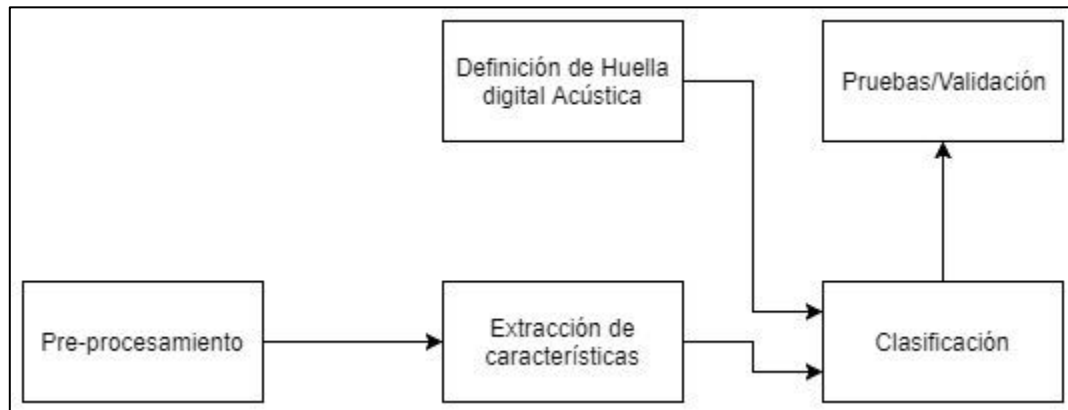


Figura 2. Diagrama de bloques del procedimiento que se buscaba desarrollar.

## 4.1 Base de datos

Para llevar a cabo el trabajo de grado se trabajó con una base de datos de una salida de campo del profesor Oscar Alberto Laverde Rodríguez de la facultad de Ciencias, departamento de Biología. La base de datos está compuesta por archivos de audio de 25 puntos de grabación, puestas a trabajar en 4 diferentes momentos. El tamaño de la base de datos es de alrededor de 700GB teniendo los datos en formato WAC. Para trabajar en el programa Audacity y en el programa MATLAB, la base de datos necesitaba ser convertida a formato WAV, por lo tanto, se convirtieron todos los archivos a WAV por medio del programa Kaleidoscope pro. Al hacer la conversión de los archivos la base de datos aumentó de tamaño a cerca de 800GB.

La base de datos consta de dos tipos de bosques en el Magdalena Medio Colombiano, uno inundable y otro no inundable. La mayoría de los datos pertenecen a bosques no inundables. La clase de cada grupo de archivos puede ser consultada en el Anexo 1. Los archivos de audio corresponden a grabadoras que se dejaron en diferentes puntos de bosques en el Magdalena Medio Colombiano. Cada grabadora tomaba muestras de 5 minutos cada 10 minutos durante el día. La cantidad de muestras tomados en un día será entonces de 144.

$$\frac{60 \text{ min}}{\text{muestras cada } 10 \text{ min}} = 6 \text{ muestras/hora} \quad (6)$$

$$24 \text{ horas} * \frac{6 \text{ muestras}}{\text{hora}} = 144 \text{ muestras} \quad (7)$$

Se debe tener en cuenta que no es constante la cantidad de audios grabados en cada punto de grabación, las carpetas no son similares en cuanto a tamaño. La posición exacta geográficamente de las grabadoras que tomaron datos se encuentra en el Anexo 2.

## 4.2 Pre-procesamiento

Para tener una idea de las características acústicas de las dos clases de bosques se ingresaron muestras de las clases a Audacity, para así analizar los sonidos en cada archivo y tratar de reconocer lo que estaba ocurriendo. Los audios difieren todos entre sí en cuanto a magnitud se trata, es por eso que se debieron normalizar todos los audios para que, al final, tuvieran la misma magnitud. Este procedimiento se llevó a cabo en Audacity y se realiza con el fin de tener una misma referencia para todos los archivos que van a analizarse.

Aparte, para asociar de manera correcta los sonidos de cada archivo, se obtuvieron los espectrogramas de esos archivos para revisar qué frecuencia tenían ciertos sonidos, cuanto tiempo duraban y cada cuanto ocurrían. Así como lo muestra la Figura 3, se muestra el espectrograma con la frecuencia en el eje vertical, el tiempo en el eje horizontal y la amplitud en el mapa de colores de la figura. Se observan bandas en el tiempo continuo entre 5kHz y 9kHz. También, se aprecia un sonido repetitivo entre 13kHz y 20kHz.

Algunas muestras de los espectrogramas obtenidos pueden consultarse en el Anexo 3. Al comparar los espectrogramas de las dos clases se encuentra que no tienen una característica particular que permita la identificación a simple vista de cada clase. Tienen comportamientos similares en cuanto a la información que el espectrograma nos puede mostrar (bandas de sonido a las mismas frecuencias, sonidos periódicos a las mismas frecuencias, etc.), por lo tanto, no se podría basar en esta característica para una clasificación.

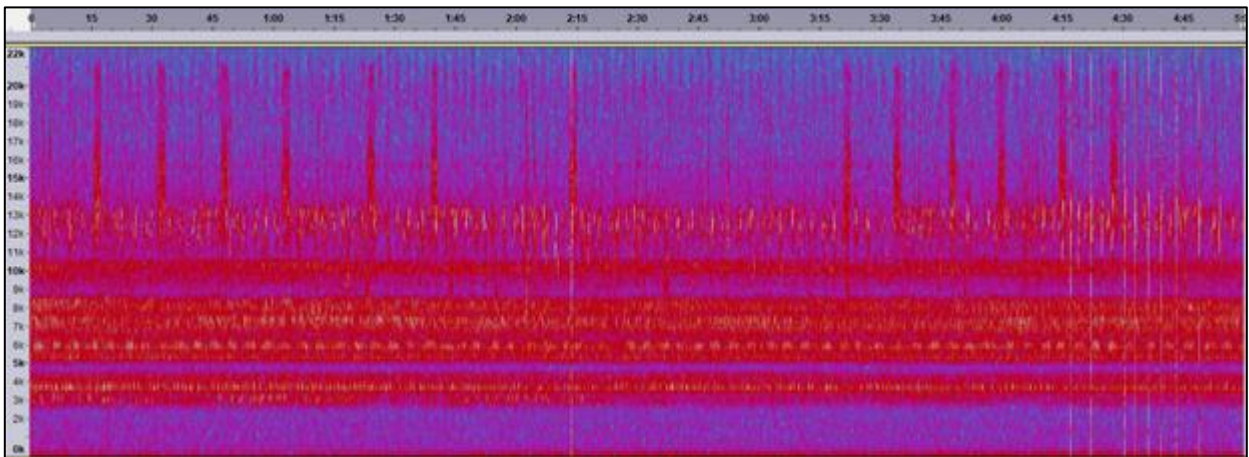


Figura 3. Espectrograma de un archivo de audio.

Para facilitar el ingreso a la base de datos, se organizaron los archivos de cada punto de grabación en carpetas por días, ya que, en la mayoría de puntos, las grabadoras tomaron datos por más de un día. Esto facilita el trabajo en la base al utilizar software de programación (R o MATLAB). Para la facilidad de los cálculos y pruebas, se organizó la base de datos en una estructura de MATLAB organizada como lo muestra la Figura 4.

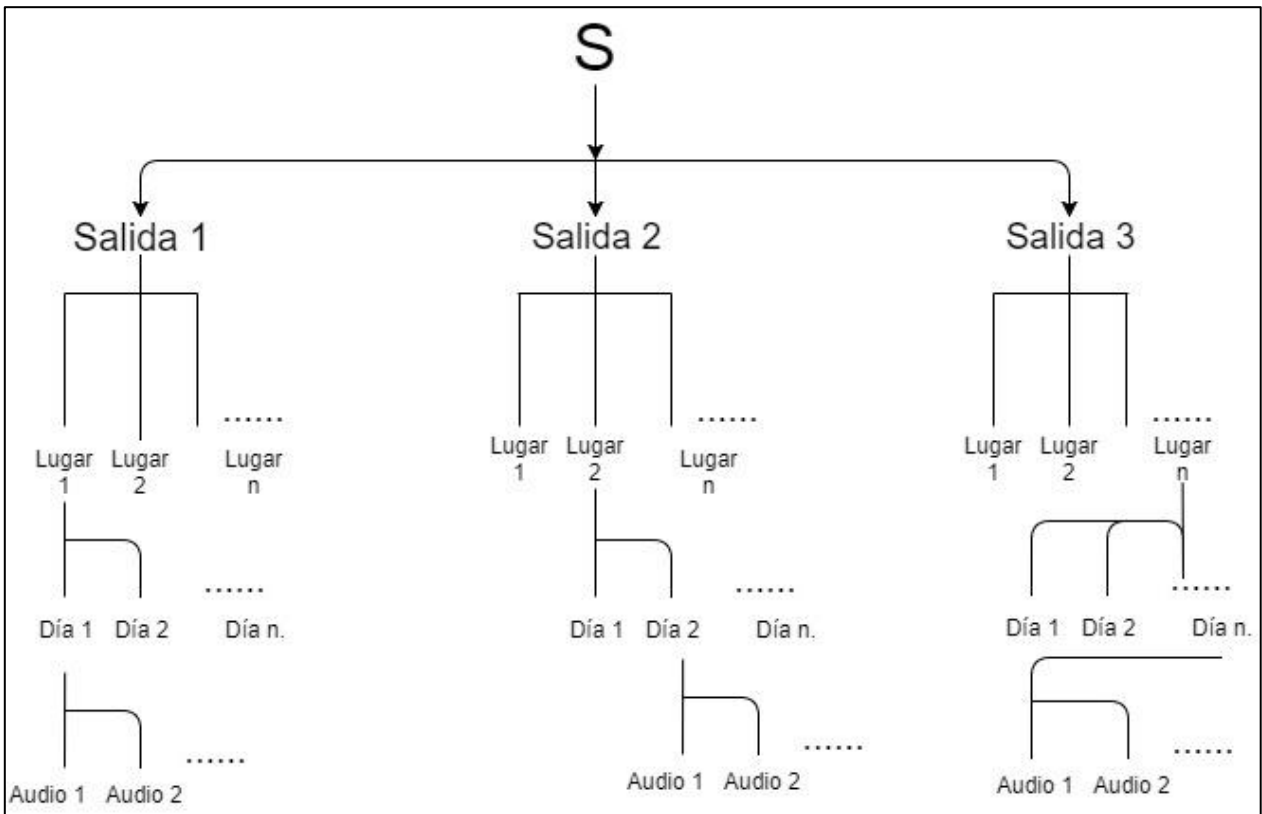


Figura 4. Estructura de MATLAB de la base de datos.

### 4.3 Extracción de características

#### 4.3.1 Trabajo en R

Se buscó, en R, obtener características de los audios a partir de un índice acústico y el cálculo de entropía de los audios, con el fin de diferenciar las dos clases de paisajes. Aunque se obtuvieron los resultados del cálculo del índice y de la entropía, no se buscaron, en los resultados, características que diferenciaran las dos clases de paisajes, ya que al ver el código fuente del índice calculado, este trabajaba con entropía, cálculo que se puede hacer por medio de MATLAB. El índice que se obtuvo de los audios fue el de Diversidad acústica del paquete de *Soundecology* (Villanueva-Rivera & Pijanowski, 2016).

#### 4.3.2 Remodelación de vectores

Se decidió trabajar por medio de MATLAB, y, al obtener los datos de cada archivo de audio, se encontró que cada audio estaba compuesto por un vector de 13 millones de posiciones. Cada uno de los archivos de la base de datos corresponden a archivos de 5 minutos de grabación. Ese tiempo de grabación corresponde, como se dijo anteriormente, a un vector de 13'000.000 de posiciones. Así como en el Reconocimiento de Iris, se busca caracterizar los archivos de audio en una forma rectangular, de esta forma, se realizó una remodelación del vector de características del archivo de audio a una matriz de 1000 filas y 13000 columnas.

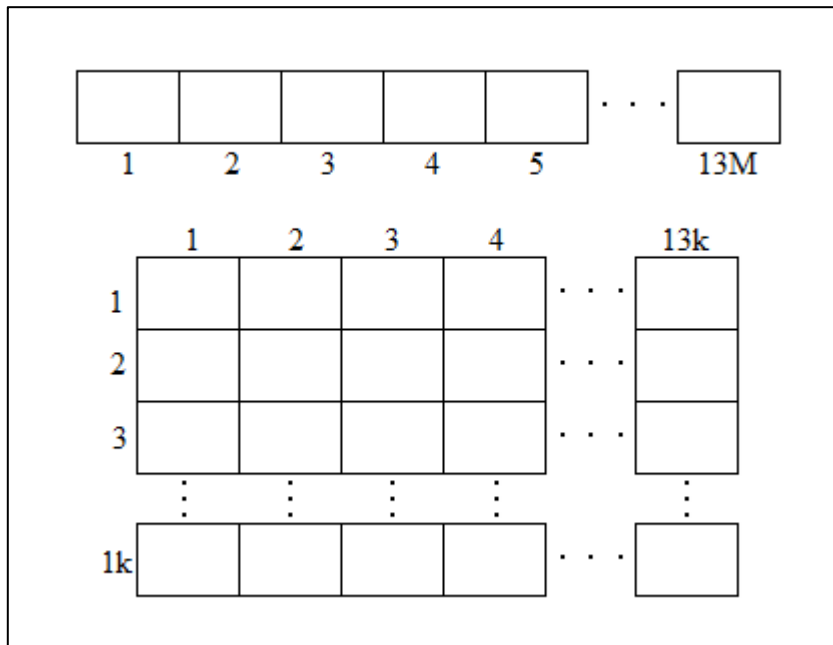


Figura 5. Representación gráfica de la remodelación del vector de características de cada archivo.

#### 4.3.3 Cepstrum

Se le aplicó el Cepstrum a los archivos de audio para encontrar qué tanto varía el espectro de frecuencias del archivo. Se presentó un problema ya que, al hacer el cálculo del Cepstrum, el espectro de frecuencias se encuentra en escala logarítmica, es por eso que, en posiciones donde el valor fuera cero, se iba a obtener un valor de infinito negativo. Esto es debido a que el logaritmo de cero es igual a infinito negativo, así como lo muestra la ecuación (8).

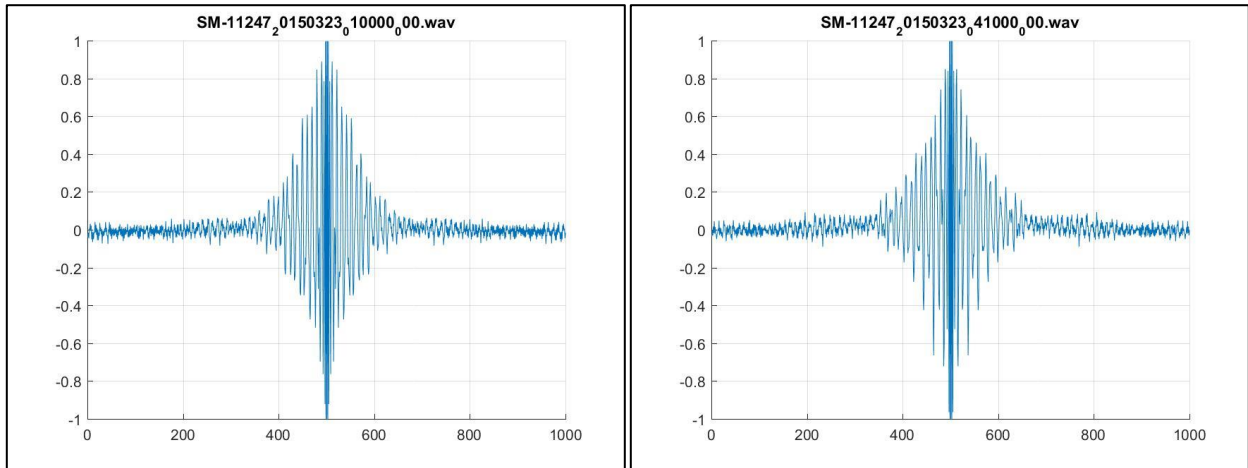
$$\log(0) = -\infty \quad (8)$$

Basados en el reconocimiento de iris, en donde se caracteriza cada subsección de información de la imagen, se busca caracterizar la matriz del cálculo del Cepstrum de cada archivo. Es por eso que, luego de aplicar el Cepstrum, se obtiene un promedio de los signos de cada columna, obteniendo así un vector de 1000 posiciones con la información del promedio de signos del Cepstrum aplicado a la matriz que contiene las características de la señal de audio.

A la hora de graficar este comportamiento, las casillas del vector donde el valor fuera infinito negativo significaban un problema ya que MATLAB no grafica un vector cuando se tienen valores de infinito o infinito negativo. Es por eso que, para obtener todas las gráficas del cálculo de la media de los signos de las columnas de las matrices a las que se les calculó el Cepstrum, se hizo una búsqueda de valores iguales a cero en la matriz inicial (antes de hacer el cálculo del Cepstrum) para reemplazar estos valores con valores cercanos a cero. Finalmente, con los valores iguales a cero cambiados, el cálculo del Cepstrum no arrojó valores iguales a infinito negativo. Este procedimiento era de suma importancia ya que, sin importar que el objetivo fuera obtener el promedio de los signos de esas casillas, al calcular el promedio de los signos de una casilla de valor infinito negativo MATLAB no realiza este cálculo, arrojando como resultado de ese promedio un valor nulo y, por lo tanto, impidiendo la representación gráfica.

Ejemplos del resultado final se pueden ver en las Figuras 6. a) y 6. b) y un ejemplo del cálculo de la media de los signos del Cepstrum de la matriz de características de los archivos de audio durante un día puede encontrarse en el Anexo 4.





6. a)

6. b)

Figura 6. a) y 6.b) Promedio de los signos del Cepstrum aplicado a una matriz con las características de un archivo de audio.

Estas gráficas, aunque mostraban un comportamiento cíclico independiente de la clase, no mostraban características válidas para establecer un criterio de clasificación. Por otro lado, este comportamiento cíclico sirvió para establecer la huella digital acústica por días.

El Cepstrum mostraba una reflexión en la posición 500 del vector, teniendo por cada vector el doble de información. Se tomaron los datos desde la posición 501 hasta la posición 1000 de cada vector.

#### 4.4 Definición de Huella Digital Acústica

Con la idea que los sonidos en los paisajes son cíclicos con periodo de un día, se pretendía conformar la huella digital acústica de cada día de grabación.

La huella digital acústica diaria se armó al tener el promedio de los signos del Cepstrum de cada archivo de audio, siendo este promedio un vector de 500 posiciones.

$$\text{Media}(\text{signo}(F^{-1}\{\log(|F\{x[t]\}|)\})) \quad (9)$$

Teniendo en cuenta que cada uno de esos vectores representa un archivo de 5 minutos de duración, obtenido cada 10 minutos durante un día (por la división de los archivos por día), se identificó que, para conformar la huella acústica diaria y basados en los archivos, se necesitan 144 vectores, ya que durante cada hora se grabó 6 veces y el día tiene 24 horas. Por lo tanto, para la conformación de la huella digital acústica diaria se formaron vectores de 72000 datos al concatenar los 144 vectores que caracterizaban cada archivo de cada día (144 vectores de 500 posiciones).

El procedimiento llevado a cabo en los numerales 4.3.4 y 4.4 se puede observar en los diagramas que se encuentran en el Anexo 5 y en el Anexo 6. El código en donde se hicieron los cálculos del promedio de los signos del Cepstrum de cada archivo y la conformación de la huella digital acústica diaria se puede encontrar en el Anexo 7.

##### 4.4.1 Interpolación de datos

Ya que la base de datos, para muchos días de grabación, se encontraba incompleta (no se tenían las grabaciones que conforman un día completo), se optó por hacer interpolación de datos a las matrices que forman la huella digital acústica de cada día pero que no tenían archivos para conformar el día completo. El código donde se hace cálculo de las horas que hacen falta en cada día y el código donde se hace la interpolación se pueden encontrar en el Anexo 8. Se realizó una interpolación con el método de *Splines* por medio de MATLAB.

#### 4.4.2 Distancia Euclidiana

Para obtener un criterio de clasificación claro se optó por calcular la distancia euclidiana, de la media de los signos del Cepstrum calculado anteriormente, entre días y así comparar, por clase, la huella de dos días iguales en un lugar igual, la huella de dos días diferentes en un lugar igual y la huella de diferentes días en lugares diferentes. Para realizar esa comparación se tomó un día de referencia por cada clase. Se obtuvo, de esas comparaciones, una matriz de 144 filas y 144 columnas. Ejemplos de estas comparaciones se pueden ver en las Figuras 7.

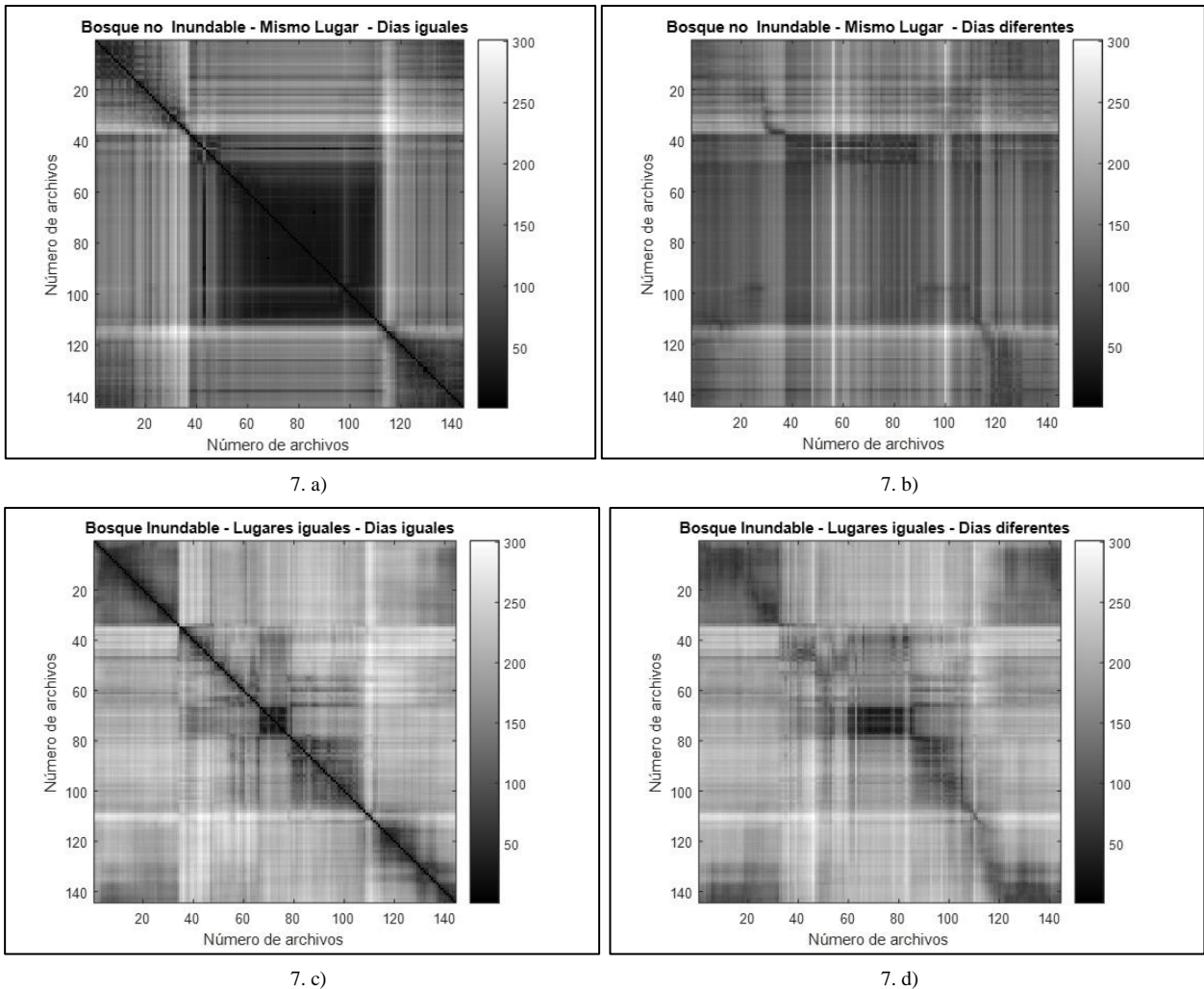


Figura 7. Ejemplos de comparaciones de Huella Digital Acústica. 7.a) Bosque no inundable, lugar igual, día igual. 7. b) Bosque no inundable, lugar igual, día diferente. 7.c) Bosque inundable, lugar igual, día igual. 7. d) Bosque inundable, lugar igual, día diferente.

Como se puede ver en los resultados, no se encontraron características que diferencien las dos clases, por lo tanto, no se tiene un criterio de clasificación claro. La información más relevante encontrada al sacar la distancia euclidiana de las huellas de una misma clase fue que tenían comportamientos similares al acercarse a las 6 de la mañana y a las 7 de la noche (color más opaco). Las horas, para caracterizar estas imágenes se pueden obtener fácilmente por medio de una regla de 3 sencilla como lo muestra la ecuación (9).

$$\frac{24 \text{ horas}}{144 \text{ archivos}} * \# \text{ de archivo} = \text{Hora del archivo} \quad (10)$$

Finalmente, se realizaron comparaciones, por medio del cálculo de la distancia euclidiana, de huellas digitales diarias pertenecientes a clases diferentes. Resultados de estas comparaciones pueden encontrarse en el Anexo 9.

#### 4.4.3 Comparación de Huellas digitales acústicas diarias entre clases.

Teniendo ya la representación de la huella diaria en vectores, se procedió a conformar una matriz para analizar las huellas digitales diarias y compararlas frente a las huellas diarias de la otra clase. Se conformaron entonces las matrices por clase concatenando cada uno de los vectores que representan las huellas digitales acústicas diarias de cada clase. Por lo que la clase 1 de paisajes, correspondiente a los bosques no inundables, tenía 136 días completos, se conformó una matriz de 136 filas y 72000 columnas. Por otro lado, como la clase 2 de paisajes, correspondiente a los bosques inundables tenía 37 días completos, se conformó una matriz de 37 filas y 72000 columnas. Estas matrices corresponden a las huellas digitales diarias de las dos clases de bosques de la base de datos.

Buscar conformar matrices de las huellas digitales diarias por cada clase se realizó con el objetivo de confirmar si el comportamiento de las huellas digitales diarias de una clase se diferenciaba de las huellas digitales diarias de la otra clase. Este análisis podría hacerse al concatenar las matrices de ambas clases y calcular, así como se hizo en el numeral 4.4.2, la distancia euclidiana de esa matriz. La matriz obtenida al concatenar las dos matrices de clases fue una gran matriz de 173 filas y 72000 columnas.

Se esperaría entonces que, si las huellas digitales diarias de cada clase son semejantes, la representación gráfica del cálculo de la distancia euclidiana de esa gran matriz de 173 filas y 72000 columnas se comportara como lo muestra la Figura 8. Esto debido a que las características, cuando son similares, tienen valores cercanos a cero (color negro) y cuando son totalmente diferentes, valores cercanos a 300 (color blanco).

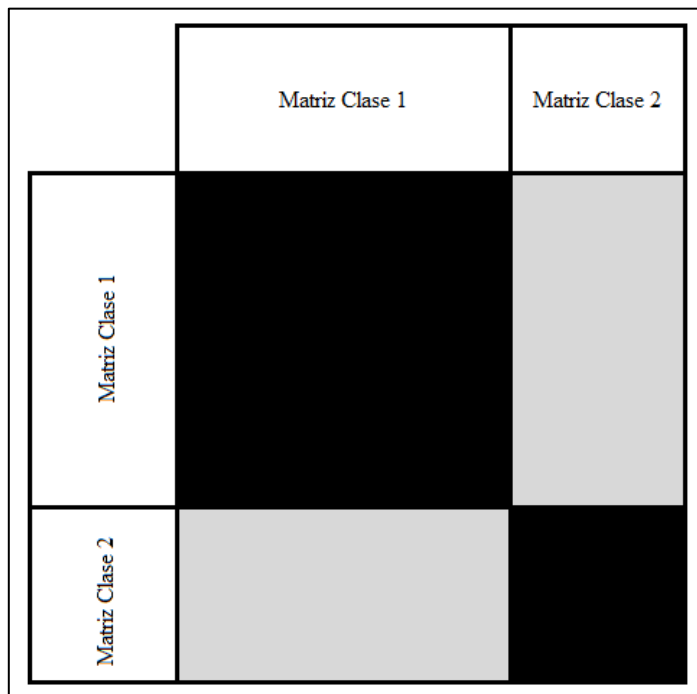


Figura 8. Expectativa la matriz del cálculo de la distancia Euclidiana entre la matriz de clases y ella misma.

Al realizar esta medida de la distancia euclidiana entre la matriz de clases y ella misma se obtuvo la imagen mostrada en la Figura 9.

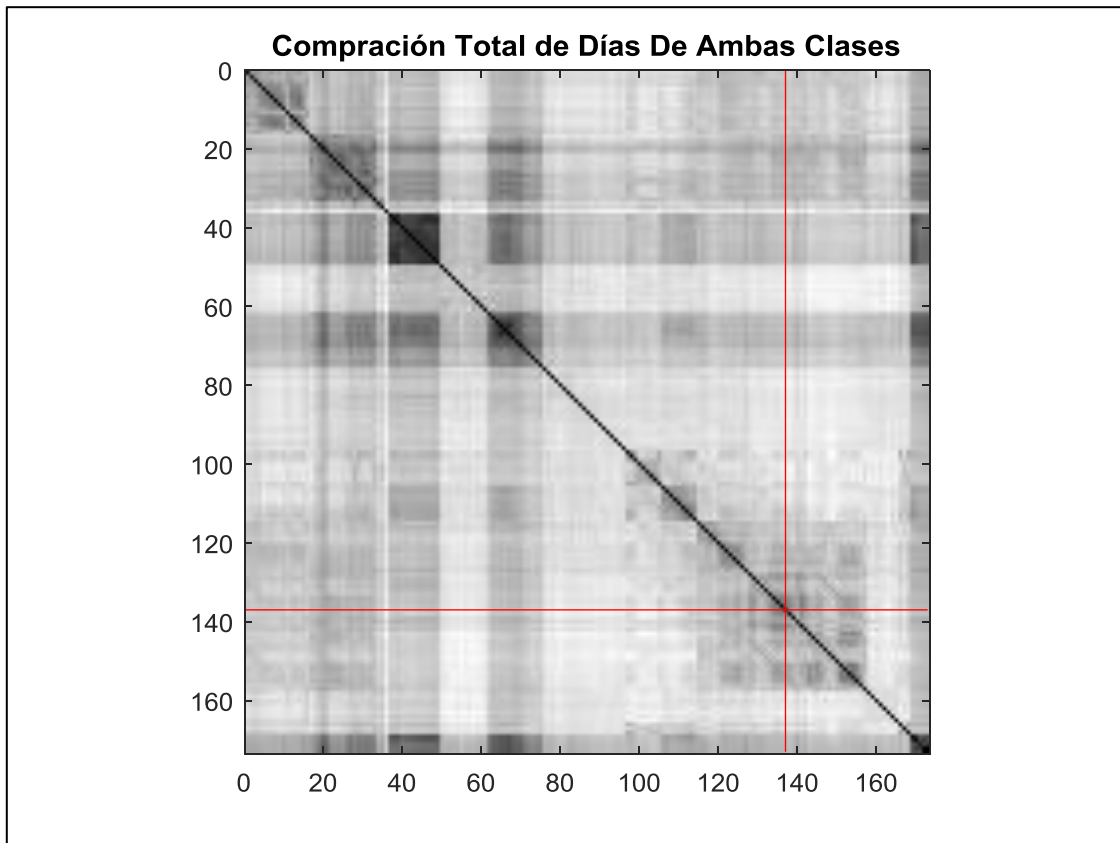


Figura 9. Matriz del cálculo de la distancia euclidiana entre la matriz de clases y ella misma.

Ya que los resultados obtenidos al hacer este cálculo no fueron similares a los esperados, se buscó la forma de encontrar la semejanza de comportamiento de las huellas digitales diarias al separarlas por Salidas de la base de datos y concatenarlas en una nueva matriz de 173 filas y 72000 columnas. Esta nueva matriz está ahora compuesta por la concatenación de los vectores correspondientes a las huellas digitales acústicas diarias ordenadas por Salida de la base de datos. Esto, para importancia del nuevo cálculo es importante ya que las dos clases de paisajes comparten archivos en ciertas salidas. El resultado de este nuevo cálculo de distancia euclidiana se puede ver en la Figura 10.

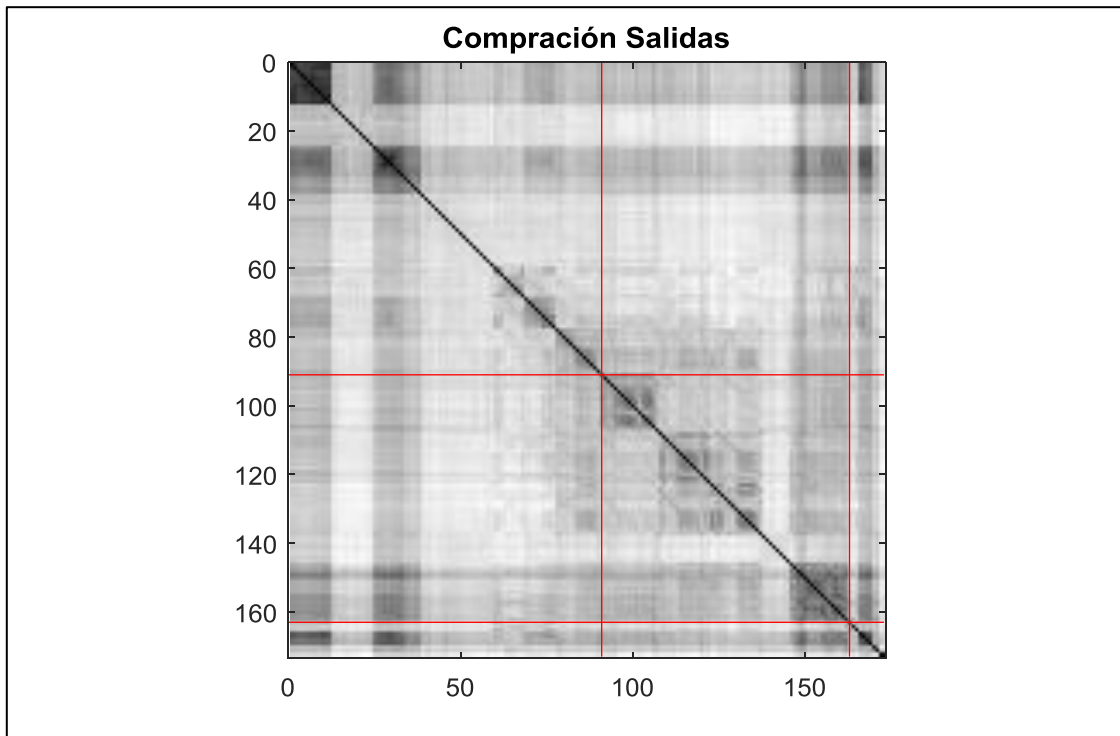


Figura 10. Matriz del cálculo de la distancia euclidiana entre la matriz de salidas y ella misma.

El código en donde se realizó este procedimiento se puede encontrar en el Anexo 10.

## 4.5 Clasificación

### 4.5.1 Máquinas de vectores de soporte

Ya que no se obtuvieron criterios de clasificación válidos, se implementó el método de Máquinas de vectores de soporte. Esto con el fin de encontrar una separación de las dos clases de paisajes que se han venido trabajando. Se utilizó un kernel lineal para separar las dos clases de bosques. Ya que la diferencia en la cantidad de archivos entre una clase y la otra es muy grande, no se puede entrenar este modelo con la totalidad de los datos de la base de datos puesto que el clasificador puede realizar una clasificación parcializada. Es por eso que se disminuyó el tamaño de la matriz de clase 1 (bosques no inundables) para que coincidiera con el tamaño de la matriz de clase 2 (bosques inundables). Se obtuvieron dos matrices de 37 filas y 72000 columnas.

Para el entrenamiento de este clasificador se tomó el 80% de los datos de cada matriz, pensando en futuras pruebas de clasificación con el 20% restante. Para no tener pérdidas de información, los datos tomados para la clasificación se obtuvieron al disminuir el número de filas y mantener el número de columnas. De esta forma, la matriz para la clasificación quedó de 29 filas y 72000 columnas. La matriz para la comprobación del método tendrá entonces un tamaño de 8 filas y 72000 columnas.

Para tener datos más acertados en la clasificación, se entrenó el clasificador por 20 veces, haciendo que, entre cada entrenamiento, las filas de la matriz ingresada entraran ordenadas aleatoriamente, haciendo que las entradas en cada entrenamiento fueran diferentes. Para obtener datos de la clasificación, se compararon los porcentajes de exactitud de cada entrenamiento, haciendo que, cuando este porcentaje sea mayor al porcentaje obtenido en el anterior entrenamiento, se guarde la información del entrenamiento de este clasificador.

Los resultados obtenidos de esta clasificación arrojaron un porcentaje promedio de exactitud del 95% para cada entrenamiento realizado. Aunque el porcentaje de exactitud de la clasificación es alto, el mapa de separación de clases mostraba algo opuesto, ya que las dos clases se superponían, así como se muestra en el Anexo 11. Los resultados obtenidos de esta clasificación se pueden ver en el Anexo 12. El código en el que se hizo la clasificación por medio de SVM se encuentra en el Anexo 13.

#### 4.5.2 Red de reconocimiento de patrones (patternet)

Por la cantidad de características que tiene cada clase, no fue posible usar el algoritmo de “trainlm” que es el indicado para un entrenamiento supervisado, que es el tipo de entrenamiento que se realizó. Es por eso que se optó por aplicar el patternet.

Al igual que la clasificación anterior (SVM), para no tener una clasificación parcializada, se disminuyó el tamaño de la matriz correspondiente a la clase 1 (bosques no inundables). Para el entrenamiento de la red neuronal también se disminuyó el tamaño de cada matriz para entrenar con el 80% de los datos y realizar pruebas de la clasificación con el 20% restante. Para el entrenamiento de esta red se utilizó la aplicación de redes neuronales de MATLAB (nnstart), en esta aplicación se seleccionó la aplicación de reconocimiento de patrones, dirigiendo así la clasificación a realizarse en una red neuronal patternet. Se escogieron 10 neuronas en la capa oculta y una neurona en la capa de salida. Se realizó un único entrenamiento con 39 iteraciones debido a la cantidad de datos ingresados.

Por los resultados obtenidos en la clasificación anterior, donde no se pudo establecer una distancia entre las dos clases, se buscaba comprobar que la clasificación de los datos no tuviera gran porcentaje de exactitud. Se tienen 4 matrices de confusión, La primera es de entrenamiento, la segunda es de validación, estas dos matrices se hallan con datos que la red neuronal conoce. La matriz de confusión es de test que se halla con datos que no conoce la red neuronal, la última es una combinación de todas las matrices anteriormente nombradas. La matriz de la figura 12 muestra el test que se hizo con datos que no conoce la red neuronal de la clase 1. Aunque la matriz de confusión arroja un entrenamiento, una clasificación y unas pruebas con exactitud igual al 100%, se comprueba que la clasificación no es correcta ya que, al hacer una verificación con los datos de la clase 1 restantes, se obtiene un error del 51.5%. Estos resultados se pueden ver en la Figuras 11 y 12.



Figura 11. Matriz de confusión del clasificador.

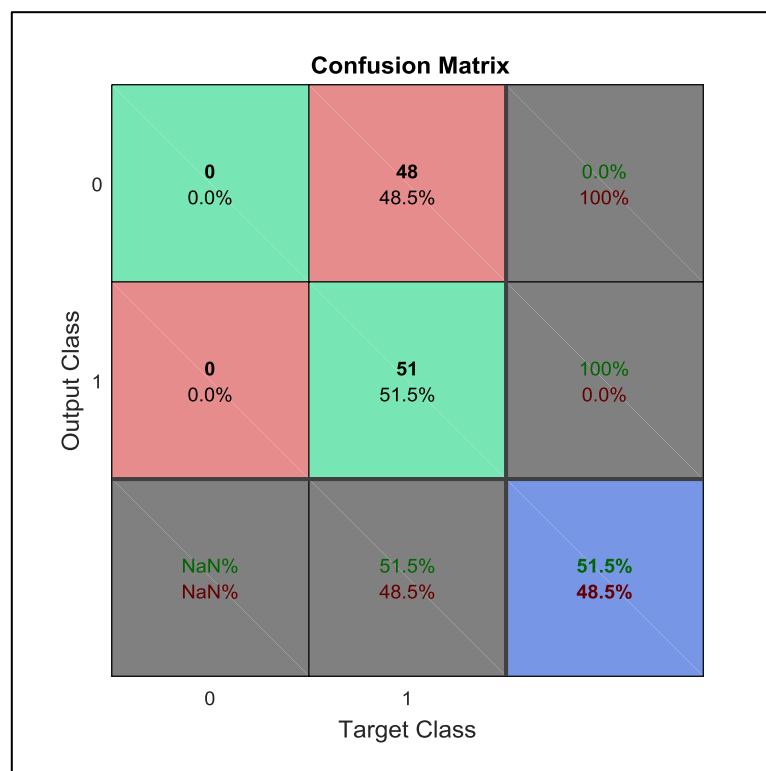


Figura 12. Matriz de confusión de verificación.

El código en el que se realizó la clasificación en esta red neuronal se encuentra en el Anexo 14.

El procedimiento que se llevó a cabo para este trabajo de grado es como lo describe el diagrama de bloques de la Figura 13.

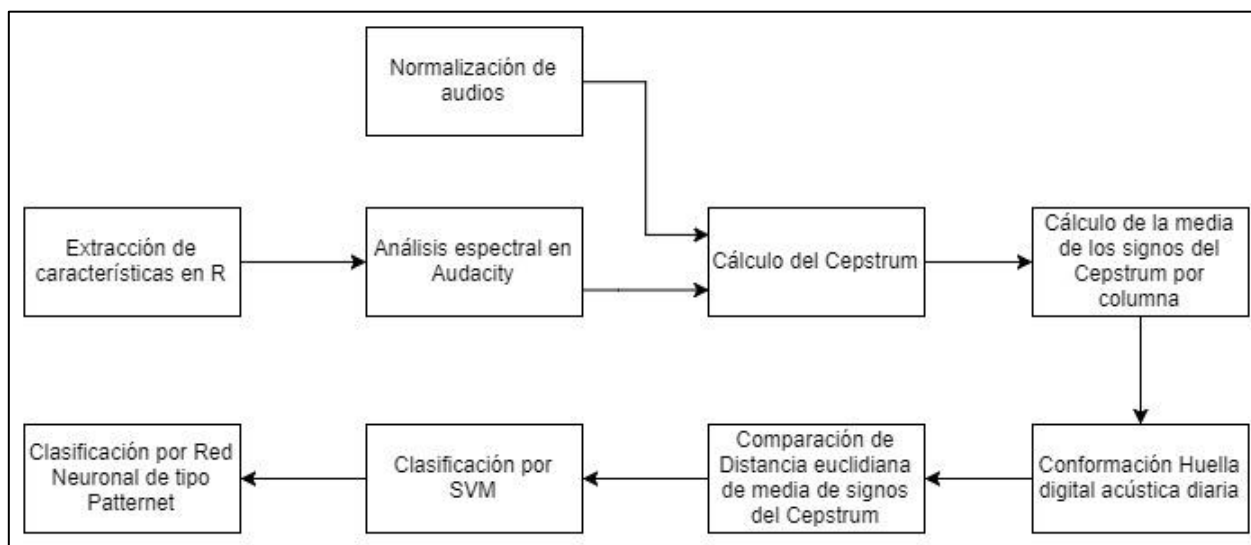


Figura 13. Diagrama de bloques de procedimiento realizado.



# CAPITULO 5: PROTOCOLO DE PRUEBAS

## 5.1 Distancia Euclidiana

Para obtener una distancia entre clases, se calculó la distancia euclidiana de las matrices de las huellas digitales acústicas diarias. entre días iguales y lugares diferentes, entre días diferentes y lugares iguales y entre días iguales y lugares iguales para tener una idea del comportamiento de los paisajes de cada clase. Finalmente, se calcula la distancia euclidiana entre varias huellas digitales diarias de diferentes clases, para encontrar características que den una idea del comportamiento de los paisajes, por día, frente a paisajes de otra clase.

Al no encontrar diferencias significativas de los comportamientos de los paisajes entre clases y al encontrar que el comportamiento del paisaje dentro de la misma clase no era semejante, no se pudo obtener un criterio para la clasificación.

Buscando un criterio para clasificación, se conformaron las matrices de las huellas de clases. Se forma esta matriz concatenando las huellas acústicas de cada uno de los días de cada clase. Se conforma una gran matriz de la huella total a partir de la concatenación de la matriz de la huella acústica de la clase 1 con la matriz de huella acústica de la clase 2. Este proceso se hace con el propósito de hacer una comparación con ella misma a partir de la distancia euclidiana, pensando en que se obtendrá información de la semejanza y diferencia de las huellas diarias de una clase comparadas con las huellas pertenecientes a la otra clase.

## 5.2 Máquinas de vectores de soporte

Se diseñó el entrenamiento como se indicó en el numeral 4.5.1. Para la verificación de la clasificación se tomaron los datos no utilizados para el entrenamiento. Todo esto con el fin de encontrar el hiperplano que divida las dos clases en una distancia separable. Se obtuvo la matriz de confusión para obtener datos de exactitud, sensibilidad y especificidad. Se obtuvo el mapa de ubicación de elementos de las dos clases.

## 5.3 Red de reconocimiento de patrones (patternet)

Se diseñó el entrenamiento de la red neuronal como se indicó en el numeral 4.5.2. Para la verificación de la clasificación de la red neuronal se tomaron los datos no utilizados para el entrenamiento. Esto se hizo con el fin de clasificar los archivos de la base de datos entre las dos clases que se tienen. Se obtuvo la matriz de confusión para obtener datos de exactitud en el entrenamiento y en las pruebas.

# CAPITULO 6: ANÁLISIS DE RESULTADOS

Debido a la gran diferencia en cantidad de archivos entre las dos clases, es mucho más complicado encontrar un criterio de separación de clases. Esta gran diferencia en cantidad, aparte de generar la dificultad para encontrar una separación de clases, genera dificultades en la etapa de entrenamiento y pruebas al hacer la clasificación.

## 6.1 Distancia Euclidiana

Los resultados arrojaron ciertos comportamientos. Uno de ellos fue que, para ciertos días, se encuentra un patrón definido y cíclico, sin embargo, muy pocos lugares cumplen con este comportamiento. Dos lugares cíclicos, al calcularles la distancia euclidiana entre ellos, no siempre tienen un comportamiento cíclico, lo que significa que los ciclos de un lugar no coinciden en la frecuencia con los ciclos que tiene el otro lugar.

Por otro lado, otro comportamiento encontrado al analizar un lugar cíclico es que tiene comportamientos similares al llegar al amanecer y al atardecer, lo que quiere decir que, llegando a esas franjas del día, los cambios en el paisaje no son apreciables en la frecuencia. Al hacer el cálculo de la distancia euclidiana entre ciertos días con día o lugar diferente, no se obtiene un comportamiento a analizar, no se encuentran un patrón definido para hacerle un análisis.

Cuando las imágenes del cálculo de la distancia euclidiana entregan imágenes con colores muy claros quiere decir que el cambio en frecuencia entre los días analizados es muy diferente entre uno y otro, lo que significa que son días con comportamientos opuestos.

En cuanto a la distancia euclidiana de la matriz de clases, se encontró que no hay coincidencias en la actividad en frecuencia de huellas de la misma clase. Al analizar la distancia euclidiana entre las salidas tampoco se encontró un comportamiento similar en frecuencia.

## 6.2 Máquinas de vectores de soporte

Aunque la clasificación fue exitosa, el mapa de las clases no arroja una separación apreciable. Esto podría estar ocurriendo debido a que todos los archivos comparten características entregadas por el Río Magdalena. También, esto ocurre gracias a la falta de archivos de algunos días y lugares de la base de datos, ya que, al aplicar la interpolación de datos, se crearon características aproximadas de los paisajes analizados para completar la base de datos y hacer un mejor análisis. Estas características agregadas en la interpolación pueden ser compartidas por las clases, agregándole características que no son propias de su clase.

Analizando la gráfica del mapa de clases, como se superponen las clases, se puede decir que la base de datos si es separable, pero por otro tipo de etiquetas diferente a la inundabilidad del paisaje. Se podrían separar las clases entre la diagonal de la gráfica y el resto de datos.

## 6.3 Red de reconocimiento de patrones (patternet)

Se obtiene una clasificación con gran exactitud por la cantidad de características que tiene cada clase.

Como se esperaba, por los resultados obtenidos al aplicar el SVM y al analizar la distancia euclidiana de la matriz de clases, la base de datos no es clasificable para esas etiquetas. Se llega a este análisis al obtener bajos índices de error en el entrenamiento, pero un alto índice de error de comprobación respecto a la prueba.

## **CAPITULO 7: CONCLUSIONES Y RECOMENDACIONES**

Observando la ubicación de las grabadoras usadas para la obtención de la base de datos, se encuentra que todas las grabaciones se realizaron cerca de una gran fuente de ruido tal como es el río Magdalena. Esto se convierte en un problema ya que les da características similares a las dos clases de audios analizadas.

Al hacer la interpolación a las matrices que representan la huella digital acústica diaria, se están agregando características aproximadas a un día (completar el número de archivos que constituye un día), y, se está agregando mucha información poco precisa a la base de datos.

Se realizó un correcto procedimiento al obtener las características de los audios en el dominio de la frecuencia, ya que esto entrega mucha más información que el análisis temporal. Sin embargo, la información irregular de la base de datos no permitió que se alcanzara el objetivo de clasificación.

La base de datos no es clasificable por el método que se utilizó.

También, se recomienda que la base de datos tenga una cantidad de muestras similares entre las clases a analizar.

# ANEXOS

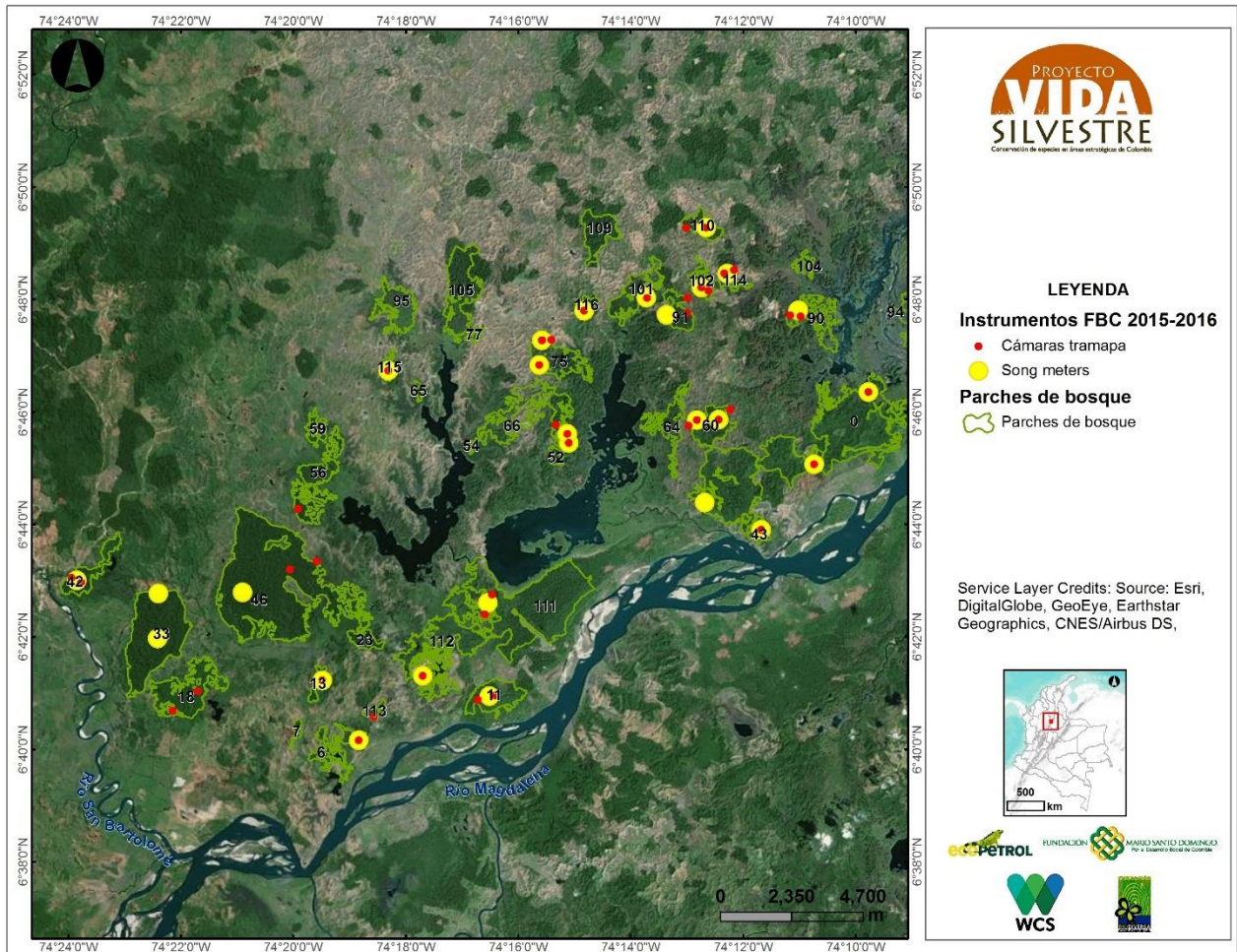
## Anexo 1

Base de datos con las etiquetas de clase para cada carpeta de audios.

[https://www.dropbox.com/sh/twk68rvq0rbsvhm/AAhR8w1--ss\\_tdgDpCoiNQKa?dl=0](https://www.dropbox.com/sh/twk68rvq0rbsvhm/AAhR8w1--ss_tdgDpCoiNQKa?dl=0)

## Anexo 2

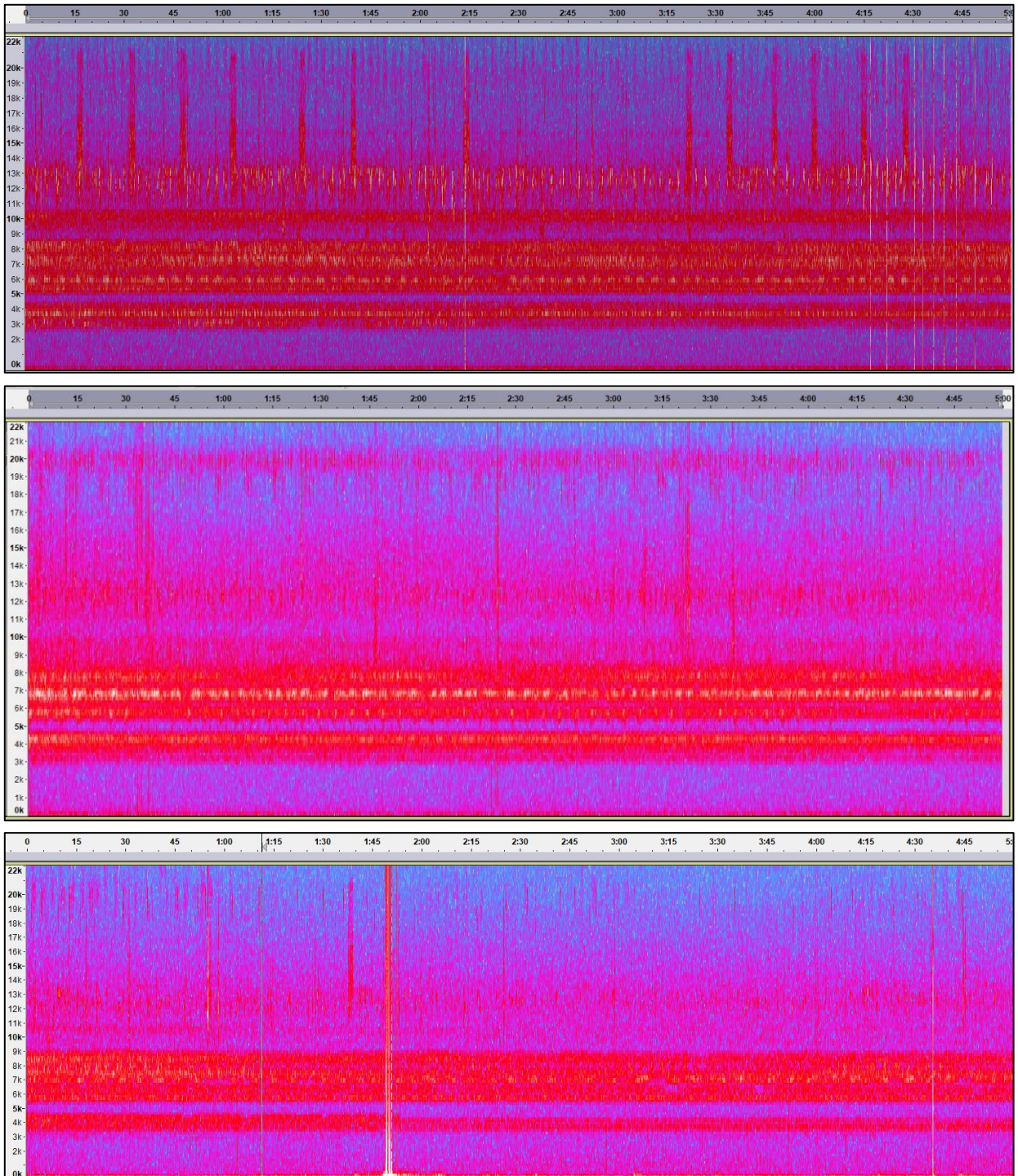
Posición geográfica de las grabadoras que tomaron los audios que conforman la base de datos.





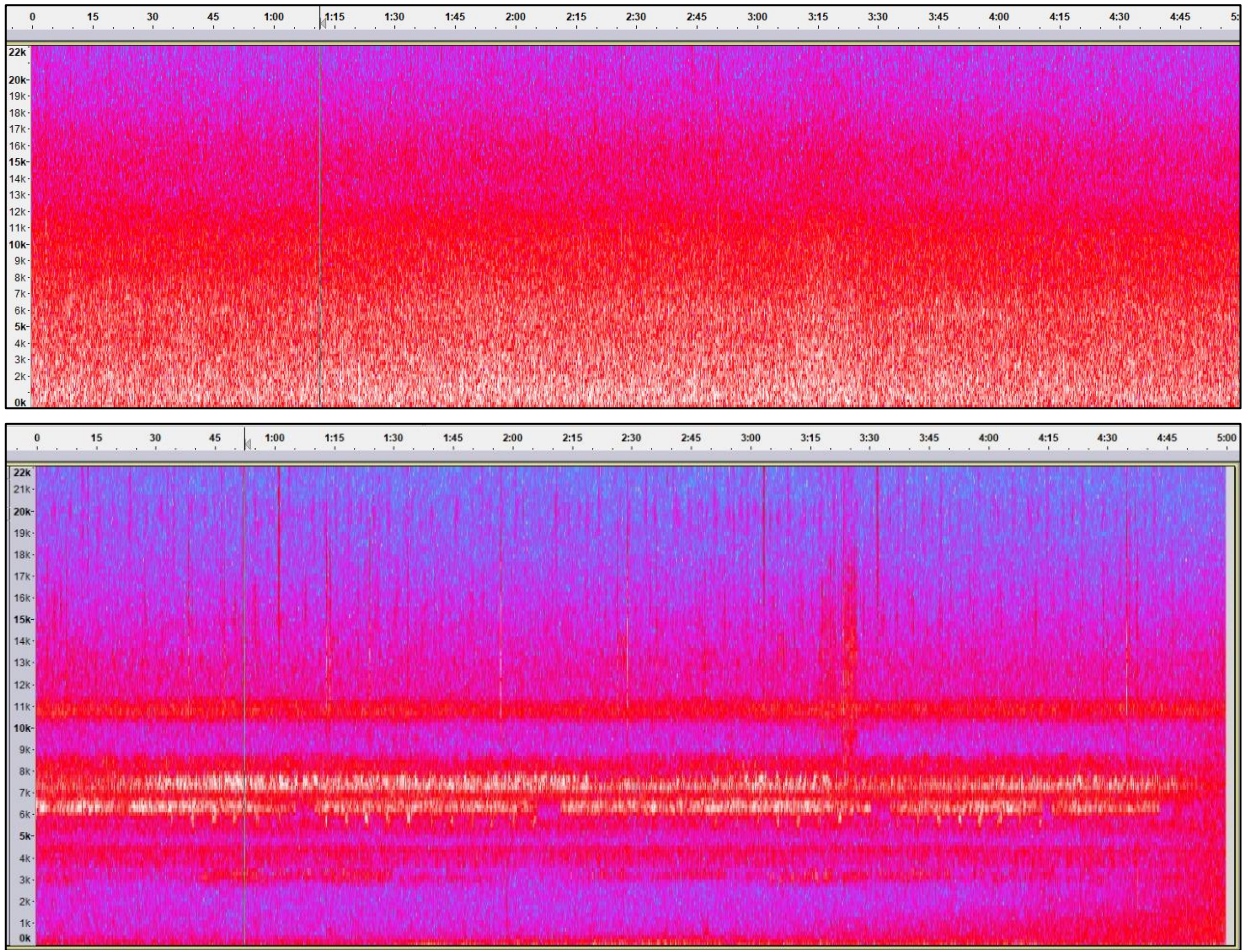
### Anexo 3

Espectrogramas de las clases de paisajes.



2. a)





2. b)

Figuras 2. a) Espectrogramas de paisaje no inundable. 2. b) Espectrogramas de paisaje inundable.

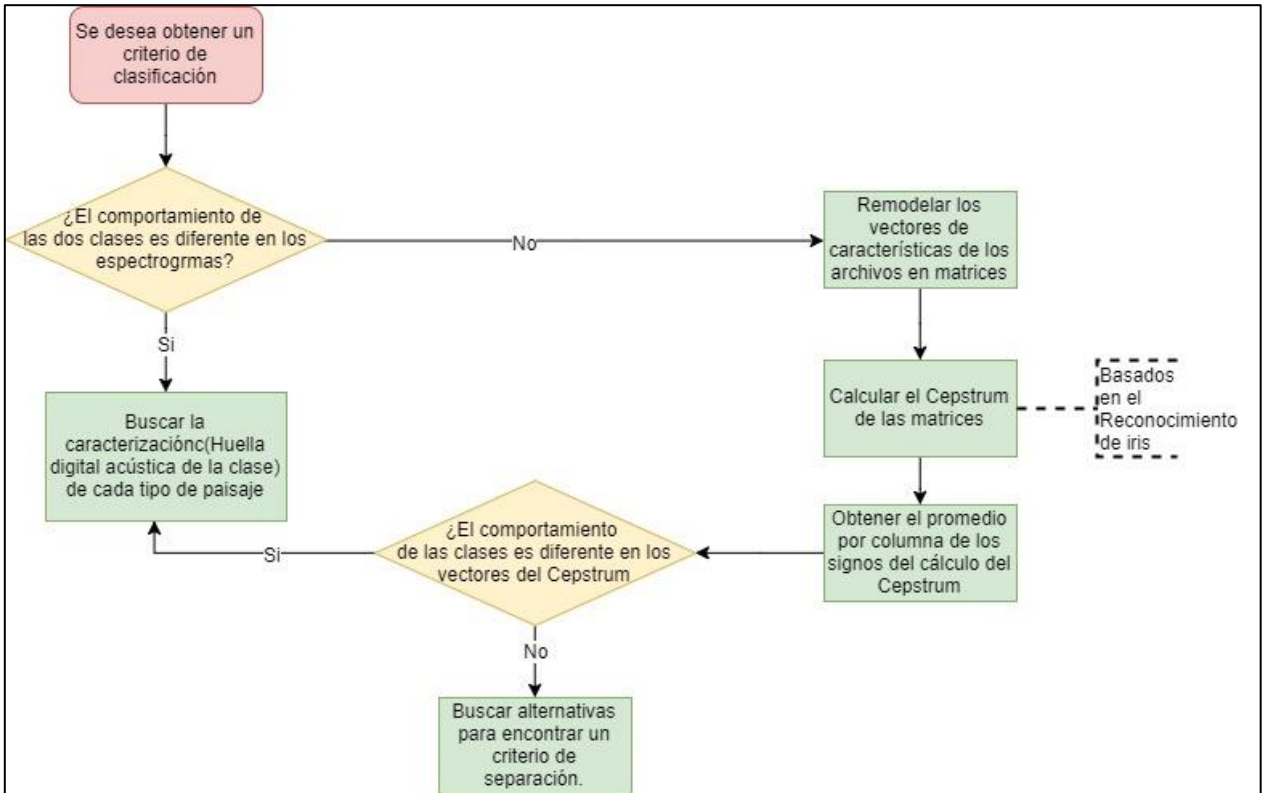
#### Anexo 4

Imágenes del promedio de los signos del Cepstrum de diferentes archivos.

[https://www.dropbox.com/sh/twk68rvg0rbsvhm/AAAhR8w1--ss\\_tdgDpCoiNQKa?dl=0](https://www.dropbox.com/sh/twk68rvg0rbsvhm/AAAhR8w1--ss_tdgDpCoiNQKa?dl=0)

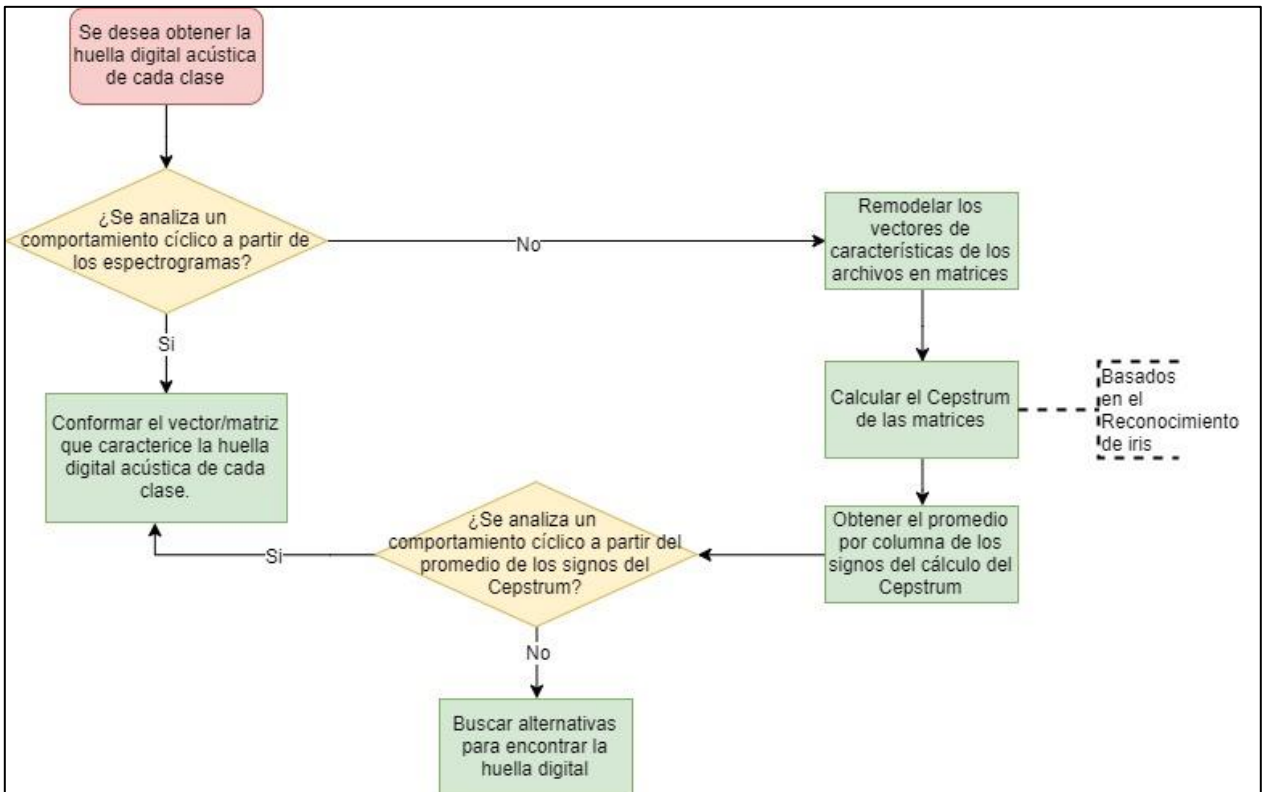
#### Anexo 5

Diagrama de flujo de los procedimientos realizados para la extracción de características en búsqueda de criterios de clasificación.



**Anexo 6**

Diagrama de flujo de los procedimientos realizados para la extracción de características en búsqueda de la Huella digital acústica de clases.



**Anexo 7**

Código para el cálculo del promedio de los signos del Cepstrum de los archivos.

[https://www.dropbox.com/sh/twk68rvq0rbsvbm/AAAhR8w1--ss\\_tdgDpCoiNQKa?dl=0](https://www.dropbox.com/sh/twk68rvq0rbsvbm/AAAhR8w1--ss_tdgDpCoiNQKa?dl=0)

**Anexo 8**

Código para el cálculo de las horas que hacen falta en cada día.

[https://www.dropbox.com/sh/twk68rvq0rbsvbm/AAAhR8w1--ss\\_tdgDpCoiNQKa?dl=0](https://www.dropbox.com/sh/twk68rvq0rbsvbm/AAAhR8w1--ss_tdgDpCoiNQKa?dl=0)

**Anexo 9**

Comparación calculada a partir de la distancia euclidiana de archivos de diferentes horas, diferentes lugares y diferentes clases. Comparación de huellas digitales acústicas diarias por medio del cálculo de distancias euclidianas para diferentes días, diferentes lugares y diferentes clases.

[https://www.dropbox.com/sh/twk68rvq0rbsvbm/AAAhR8w1--ss\\_tdgDpCoiNQKa?dl=0](https://www.dropbox.com/sh/twk68rvq0rbsvbm/AAAhR8w1--ss_tdgDpCoiNQKa?dl=0)

**Anexo 10**

Código de conformación de matriz de clases.

[https://www.dropbox.com/sh/twk68rvq0rbsvbm/AAAhR8w1--ss\\_tdgDpCoiNQKa?dl=0](https://www.dropbox.com/sh/twk68rvq0rbsvbm/AAAhR8w1--ss_tdgDpCoiNQKa?dl=0)

**Anexo 11**

Mapa de clases obtenido al aplicar SVM

[https://www.dropbox.com/sh/twk68rvq0rbsvbm/AAAhR8w1--ss\\_tdgDpCoiNQKa?dl=0](https://www.dropbox.com/sh/twk68rvq0rbsvbm/AAAhR8w1--ss_tdgDpCoiNQKa?dl=0)

**Anexo 12**

Resultados de la clasificación por medio de SVM

[https://www.dropbox.com/sh/twk68rvq0rbsvbm/AAAhR8w1--ss\\_tdgDpCoiNQKa?dl=0](https://www.dropbox.com/sh/twk68rvq0rbsvbm/AAAhR8w1--ss_tdgDpCoiNQKa?dl=0)

**Anexo 13**

Código de aplicación de SVM

[https://www.dropbox.com/sh/twk68rvq0rbsvbm/AAAhR8w1--ss\\_tdgDpCoiNQKa?dl=0](https://www.dropbox.com/sh/twk68rvq0rbsvbm/AAAhR8w1--ss_tdgDpCoiNQKa?dl=0)

**Anexo 14**

Código de clasificación por medio de la red neuronal.

[https://www.dropbox.com/sh/twk68rvq0rbsvbm/AAAhR8w1--ss\\_tdgDpCoiNQKa?dl=0](https://www.dropbox.com/sh/twk68rvq0rbsvbm/AAAhR8w1--ss_tdgDpCoiNQKa?dl=0)



# BIBLIOGRAFÍA

- Amos Gilat. (2012). *MATLAB An Introduction With Applications*. John Wiley & Sons.  
<https://doi.org/10.1007/s13398-014-0173-7.2>
- Barkat, M. (2005). Chapter 6: Parameter Estimation. In *Signal Detection and Estimation, Second Edition*.
- Bergland, G. D. (1969). A guided tour of the fast Fourier Transform. *IEEE Spectrum*, 6(7), 41–52.  
<https://doi.org/10.1109/MSPEC.1969.5213896>
- Caycedo-Rosales, P. C., Ruiz-Muñoz, J. F., & Orozco-Alzate, M. (2013). Reconocimiento automatizado de señales bioacústicas: Una revisión de métodos y aplicaciones. *Ingeniería Y Ciencia - Ing.cienc.*, 9(18), 171–195. Retrieved from <http://publicaciones.eafit.edu.co/index.php/ingciencia/article/view/2017>
- Fabbri, G., & Abagnali, V. (2012). *Sound Waves : Propagation, Frequencies, and Effects*.
- Haykin, S. (1994). Neural networks-A comprehensive foundation. New York: IEEE Press. Herrmann, M., Bauer, H.-U., & Der, R. <https://doi.org/10.1017/S0269888998214044>
- Huang, Y.-P. H. Y.-P., Luo, S.-W. L. S.-W., & Chen, E.-Y. C. E.-Y. (2002). An efficient iris recognition system. *Proceedings. International Conference on Machine Learning and Cybernetics*, 1(November), 4–5.  
<https://doi.org/10.1109/ICMLC.2002.1176794>
- Kasten, E. P., Gage, S. H., Fox, J., & Joo, W. (2012). The remote environmental assessment laboratory's acoustic library: An archive for studying soundscape ecology. *Ecological Informatics*, 12, 50–67.  
<https://doi.org/10.1016/j.ecoinf.2012.08.001>
- Koutroumbas, S. Theodoridis, K. (2003). *Pattern Recognition. Pattern recognition* (Vol. 8).  
<https://doi.org/10.1016/B978-012369531-4/50016-0>
- Leemis, Lawrence M. (2016). Learning Base R. *Journal of Statistical Software*, 69(Book Review 4), 8–10.  
<https://doi.org/10.18637/jss.v069.b04>
- MATLAB. (2016a). patternnet. Retrieved from  
<https://www.mathworks.com/help/nnet/ref/patternnet.html>
- MATLAB. (2016b). Support Vector Machines for Binary Classification. Retrieved from  
<https://www.mathworks.com/help/stats/support-vector-machines-for-binary-classification.html>
- Oppenheim, A. V., & Schaffer, R. W. (2004). From frequency to quefrequency: A history of the cepstrum. *IEEE Signal Processing Magazine*, 21(5). <https://doi.org/10.1109/MSP.2004.1328092>
- Pijanowski, B. C., Farina, A., Gage, S. H., Dumyahn, S. L., & Krause, B. L. (2011). What is soundscape ecology? An introduction and overview of an emerging new science. *Landscape Ecology*, 26(9), 1213–1232. <https://doi.org/10.1007/s10980-011-9600-8>
- Pijanowski, B. C., Villanueva-Rivera, L. J., Dumyahn, S. L., Farina, A., Krause, B. L., Napoletano, B. M., ... Pieretti, N. (2011). Soundscape Ecology: The Science of Sound in the Landscape. *BioScience*, 61(3), 203–216. <https://doi.org/10.1525/bio.2011.61.3.6>
- Romero, T. (2009). Hacia una definición de Sedentarismo. *Revista Chilena de Cardiología*, 28(4), 409–413.

<https://doi.org/10.4067/S0718-85602009000300014>

Schafer, R. M. (1977). *El paisaje sonoro y la afinación del mundo*. Intermedio Editores.

Sueur, J., Farina, A., Gasc, A., Pieretti, N., & Pavoine, S. (2014). Acoustic indices for biodiversity assessment and landscape investigation. *Acta Acustica United with Acustica*, 100(4), 772–781. <https://doi.org/10.3813/AAA.918757>

Sueur, J., Pavoine, S., Hamerlynck, O., & Duvail, S. (2008). Rapid acoustic survey for biodiversity appraisal. *PLoS ONE*, 3(12). <https://doi.org/10.1371/journal.pone.0004065>

Villanueva-Rivera, L. J., & Pijanowski, B. C. (2016). Soundscape Ecology. Retrieved from <http://ljevillanueva.github.io/soundecology/>